Development of digital imaging technologies for the segmentation of solar features and the extraction of filling factors from SODISM images

Amro F. A. Alasta

A Thesis Submitted for the Degree of

Doctor of Philosophy

School of Electrical Engineering and computer Science

Faculty of Engineering and Informatics

University of Bradford

October 2018

ABSTRACT

Key words:
Solar Diameter Imager And Surface Mapper (SODISM), Solar Imaging, Sunspots, Digital Imaging, Segmentation, Resolution Enhancement, Discrete Wavelet Transforms, Solar Filling Factor.

Solar images are one of the most important sources of available information on the current state and behaviour of the sun, and the PICARD satellite is one of several ground and space-based observatories dedicated to the collection of that data. The PICARD satellite hosts the Solar Diameter Imager and Surface Mapper (SODISM), a telescope aimed at continuously monitoring the Sun. It has generated a huge cache of images and other data that can be analysed and interpreted to improve the monitoring of features, such as sunspots and the prediction and diagnosis of solar activity.

In proportion to the available raw material, the little-published analysis of SODISM data has provided the impetus for this study, specifically a novel method of contributing to the development of a system to enhance, detect and segment sunspots using new hybrid methods. This research aims to yield an improved understanding of SODISM data by providing novel methods to tabulate a sunspot and filling factor (FF) catalogue, which will be useful for future forecasting activities.

The developed technologies and the findings achieved in this research will work as a corner stone to enhance the accuracy of sunspot segmentation; create efficient filling factor catalogue systems, and enhance our understanding of SODISM image enhancement. The results achieved can be summarised as follows:

i) Novel enhancement method for SODISM images.

ii) New efficient methods to segment dark regions and detect sunspots.

iii) Novel catalogue for filling factor including the number, size and sunspot location.

v) Novel statistical method to summarise FFs catalogue.

Image processing and partitioning techniques are used in this work; these methods have been applied to remove noise and detect sunspots and will provide more information such as sunspot numbers, size and filling factor. The performance of the model is compared to the fillers extracted from other satellites, such as SOHO. Also, the results were compared with the NOAA catalogue and achieved a precision of 98%. Performance measurement is also introduced and applied to verify results and evaluate proposal methods.

Algorithms, implementation, results and future work have been explained in this thesis.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgments.

The main part of this is based on six peer reviewed papers published or accepted for publication in different academic journals, posters and conferences.

PUBLICATIONS

Journal Papers

- A. Alasta, A. Gamudi, R. Qahwaji, S. Ipson, A. Hauchecorne, and M. Meftah “New method of Enhancement using Wavelet Transforms applied to SODISM Telescope” Published in Advances by Space Research, May 2018. Available at : https://www.sciencedirect.com/science/article/pii/S0273117718306112 .

- A. Alasta, A. Algamudi, , F. Almesrati, M. Meftah and R. Qahwaji, “Filling Factors of Sunspots in SODISM Images” Published by Annals of Emerging Technologies in Computing (AETiC), UK., Aprel 2019. Available at: http://aetic.theiaer.org/archive/v3/v3n2/p1.html

- Alasta, M. Meftah ,R. Qahwaji A. Algamudi and F. Almesrati, “Analysis Filling Factor Catalogue of Different Wavelength SODISM Images” Accepted in International Conference on Emerging Technologies in Computing August 2019 (iCETiC '19) London Metropolitan University, London, UK.

Conference Papers

- A. Alasta, A. Gamudi, R. Qahwaji, S. Ipson, and T. Nagim 2017. “Automatic sunspots detection on SODISM solar images”. In 2017 Seventh International Conference on Innovative Computing Technology (INTECH). IEEE, pp. 115–119.
  Available at: http://ieeexplore.ieee.org/document/8102429/

- A. Alasta, A. Gamudi, R. Qahwaji, and F. Almesrati. “Identification of Sunspots on SODISM Full-Disk Solar Images” published by IEEE International Conference on Computing, Electronics and Communications Engineering 2018 iCCECE’18 held at University of Essex, Southend, UK at Aug. 16 and 17, 2018.
  Available at: https://ieeexplore.ieee.org/document/8658432/authors [Accessed March 19].

- T. Nagem, R. Qahwaji, S. Ipson, A. Alasta. "Predicting solar flares by convert GOES X-ray data to Gramian Angular Fields (GAF) images", The World Congress on Engineering 2018 (WCE 2018), London, UK.

Posters

- R. Qahwaji, T. Colak, O. Ahmad, S. Ipson, A. Alasta, L. Velikov, and A. Gamudi “Applied image processing and machine learning research in the space weather domain” EPSRC engineering and physics research council 13 October 2014, University of Bradford, Bradford.

- A. Alasta, A. Algamudi, R. Qahwaji, S. Ipson “Resolution Enhancement using wavelet Transforms Applied to SODISM Images” in the 8th International Conference on Computer Science and Information Technology 2018, Amman 11-12, July, 2018, University of Bradford, UK.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Rami Qahwaji, for his encouragement, valuable support, and wise guidance. Professor Qahwaji has supported me in the course of my research as an older brother, and has been honest and kind with his time and ideas. Without him, I would not have been able to complete this thesis. I am always grateful for his patience and impressed by his wide knowledge of science, and I believe that his character will cast a reflection on me for a lifetime.

I could never forget to thank my second supervisor, Dr. Stan Ipson, who has been generous with his time, and has inspired me with his various ideas and has always been a source of invaluable knowledge. He has reviewed my papers and the thesis, giving valuable constructive feedback, on detail big and small.

I would like to thank Dr. Omar Ahmed, who supported me prior to his departure from Bradford University, for his help with the collection of data. His advice was also invaluable.

Finally, I must give heartfelt thanks to all my family, starting with my lovely wife and parents who have always supported and trusted me. My brothers and sisters have always been my source of love, and they have stood next to me during difficult times in my studies in the UK. Last, but not least, I should thank all of my friends and especially Abdurzak Algamudi, for their support and help.

Thank you

Amro F. Alasta

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AR	Active region
ASAP	Advanced Systems Analysis Program
CCD	Charge Coupled Device
CLV	Centre-to-Limb Variation
CME	Coronal Mass Ejection
CNES	Centre National d'Etudes Spatiales
CS	Cycle Spanning
DWT	Discrete Wavelet Transform
FAR	False Acceptance Rate
FF	Filling Factor
FITS	Flexible Image Transport System
FRR	False Rejection Rate
IDWT	Inverse Discrete Wavelet Transform
IQR	Inter Quartile Range
LR	Low Resolution
MMSE	Minimum Mean-Squared Error
MSE	Mean Squared Error
NAE	Normalized Absolute Error
NCC	Normalized Cross-Correlation
NOAA	National Oceanic and Atmospheric Administration
PPDC	PICARD Payload Data Centre
PREMOS	Precision Monitor Sensor
PSNR	Peak Signal to Noise Ratio
RS	Resolution Synthesis
SC	Structural Content
SDO	Solar Dynamics Observatory
SNR	Signal to Noise Ratio
SODISM	Solar Diameter Imager and Surface Mapper
SOHO	Solar and Heliospheric Observatory
SOVAP	Solar Variability PICARD
SS	Sunspot
SSI	Solar Spectral Irradiance

SSIM	Structural Similarity-Based Metric
SW	Space Weather
UV	Ultraviolet

CHAPTER ONE

1. Introduction

1.1 Background

The Sun is a basic energy source upon which all life on Earth, including humans, depend; however, it is not a static system as it has the potential for a massive detrimental impact on humans (Gray et al., 2010). Moreover, as humans are becoming more dependent on technology, this makes us more vulnerable, and great damage could potentially be caused by extreme solar activity associated with extreme flare and coronal mass ejection events (Plunkett, 2005). Such solar activity has occurred in the past, but could today have damaging effects on a wide range of space and ground-based human activities, including communications, navigation, electrical power distribution, and air, sea and ground transportation (N.R.C. Committee on the societal and economic impacts of and Space Weather, 2009)(Kumar, 2017). This realisation provides a significant incentive to improve our knowledge of the way the Sun works in order to be able to predict its future behaviour and the consequent effects on planet Earth.

Solar flares, shown in Figure 1, are sudden, short-lived bursts of energy on the Sun's surface, lasting from minutes to hours. They are connected to other, longer-lived solar features, including sunspots (Colak and Qahwaji, 2007). The study of these phenomena can improve prediction capabilities and, hence, the ability to apply preventative measures to avoid or reduce the influence on vulnerable industries of violent solar flares and other space weather phenomena. PICARD is one of the satellites providing data through Solar Diameter Imager and Surface Mapper (SODISM) images. It provides images at the following five wavelengths: 215 nm, 393 nm, 535 nm, 607 nm, and 782 nm (M. Meftah et al., 2014a). Additionally, investigating SODISM wavelengths can help in identifying which plane of a hyper-spectral image is best suited for the techniques outlined in this thesis.

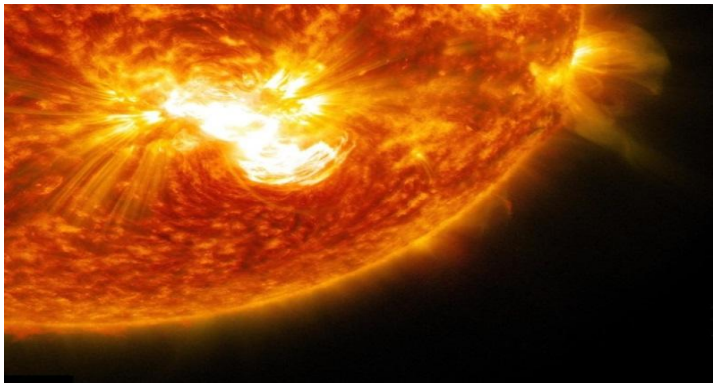


Figure 1: Huge solar flare erupts from the biggest sunspot in 24 years recorded by the NASA Solar Dynamics Observatory on Oct. 24th, 2014[1]

During the twentieth century, image-based solar data records were the main source of data for monitoring solar activities, such as sunspots and bright regions (Hathaway, 2015). These are still of interest, but the amount of data available in monitoring the activity of the Sun has increased substantially in terms of volume and variety, and this is due to the development of various ground and space-based observations (National Aeronautics and Space Administration, 2017). These observations are beneficial as they enable the application of large-scale analyses using sophisticated statistical approaches to extract and compare findings (National Aeronautics and Space Administration, 2017).

In practice, in order to study the fine structures of the Sun's surface, it is necessary to apply techniques that enhance solar images and improve the visibility of several solar features, such as sunspots and flares. Moreover, recent advances in image processing and machine learning techniques provide opportunities for the development of automatic algorithms that can be used to analyse and monitor solar activity that affect life on Earth.

In this research, the aim is to design a fully automated enhancement segmentation system that can detect solar-disk sunspots by analysing solar SODISM images; the fundamental steps are shown in Figure 2. This system uses SODISM images and applies a new technique using interpolation of the high-frequency sub-bands obtained by Discrete Wavelet Transforms (DWT), followed by applying hybrid methods for the segmentation phase, detecting

[1] https://www.space.com/27540-huge-solar-flare-from-giant-sunspot.html

sunspots, and finally, calculating the filling factor (FF) and creating novel filling factor catalogue.

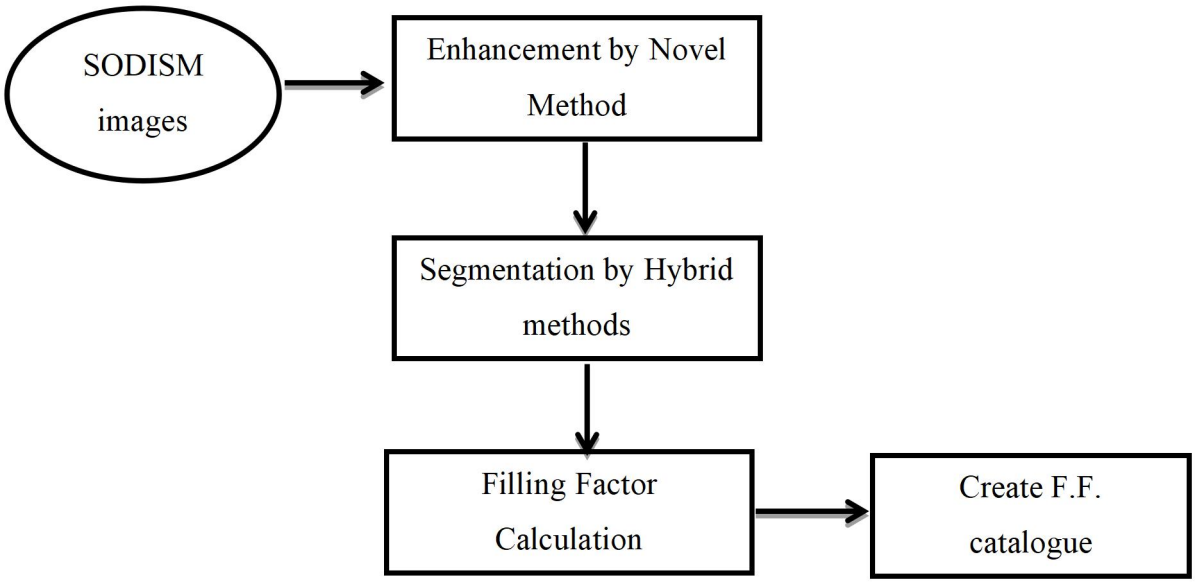


Figure 2: Block diagram of suggested system for automatic detection of sunspots.

1.2 Solar activities

1.2.1 Solar wind

The solar wind is different from wind on the Earth's surface; one million tons of hot plasma is carried by the solar wind at a temperature 5 kelvins away from the sun every second. The plasma acts like an electrically conducting fluid, carrying with it a coronal mass ejection magnetic field generated from electrical currents in the Sun's corona (NASA, 2004). Inversely, the power of the magnetic field will be decreased when distance is increased and plasma particles will escape from the sun because has it has sufficient kinetic energy. The solar wind becomes part of the Sun's corona into interplanetary space. The speed of a solar stream varies between 300 to 1000 km/s, depending on distance from the sun. Varying speeds will produce low magnetic regions and regions in which the magnetic field is amplified (NASA, 2004).

1.2.2 Sunspots

Sunspots were first observed soon after the invention of the telescope, and were observed between 1610 and 1613 by several astronomers,

including Galileo Galilei, Johann Goldsmid, Thomas Harriot, and Christoph Scheinier (‘Galileo’s Sunspot Letters’, no date). Dark features in the photosphere are known as sunspots, and these vary from simple pore-like structures to very complex structures. The spots are darker than the surrounding photosphere because they are cooler (Aeronautics, 2011). Moreover, they have very strong, complicated magnetic fields, and their sizes range from approximately 300 km, to diameters in excess of 100,000 km, a size nearly eight times the diameter of the Earth. Furthermore, the lifetime of a sunspot varies from less than one hour to more than six months (Eddy, 2010).

The surface of the Sun contains various features, such as sunspots, faculae and networks (Eddy, 2010). Sunspots do not cover the whole surface of the Sun, but rather occur in two bands about 15-20 degrees wide in latitude, which surround the Sun on either side of the solar equator. The band latitudes are not fixed but change according to the solar cycle. Observing the solar cycle on the north and south of the equator shows that sunspots appear at an average latitude of about 25-30 degrees of a minimum solar cycle. New sunspots appear closer to the equator than before as a result of the progress of the solar cycle. However, no sunspots are found at latitudes greater than approximately 70 degrees, and the final sunspots usually appear between 5-10 degrees latitude[2].

The beginning of a sunspot is usually a tiny dot, called a pore (Yu, Deng and Feng, 2014). Not all pores survive to become mature; however, those that survive will develop in pairs within the first stage of their existence. These subsequently grow in terms of size and area to reach their final stage when they then become known as sunspots. These have a dark region in the centre, called the umbra, which are usually surrounded by a light region, known as the penumbra. There are two types of spots, namely leaders and followers, and the lifetime of sunspot groups depend on whether they are leaders or followers. The leader’s life is always longer than that of the followers, and sunspot groups start to disappear slowly when the follower spots begin to lose definition and they fade away in stages. In many cases,

[2] https://starchild.gsfc.nasa.gov/docs/StarChild/questions/question17.html

the follower spots split into smaller individual spots, and continue to split into smaller units until they slowly fade away. However, the leader has a slower rate of decay and usually keeps its shape for longer, although a slight decrease in brightness may occur in the umbral and penumbral areas. This gives the leader spot several weeks of life after the rest of the spots within the group have faded away. The leader spots near the minimum of the solar cycle exist for longer than those near the maximum of the cycle (Eddy, 2010)., introduced by Waldmeier (Waldmeier M., 1976), and the second is the McIntosh classification method (McIntosh, 1990), invented by McIntosh of the US National Oceanic and Atmospheric Administration's (NOAA)[3] Space Environment Center (SEC). The NOAA sunspot catalogue records the data of many solar observatories which have been tracking sunspot regions and provide a wide reference catalogue recording their date, time, location, physical properties, magnetic classification, sunspot area and the active. There are two major types of sunspot classification; the first is the Zurich classification region number. In fact, the NOAA catalogue is the main source for comparison in this research, additional to the comparison with the SOHO filling factor results. However, the second method was introduced in 1966 and is a modification of the Zurich sunspot method (Izenman, 1983). According to McIntosh's classification, three parameters are used to classify the primary properties of a sunspot: thus, Z represents the Modified Zurich Class, p describes the penumbra of the largest spot, and c represents the sunspot description. Figure 3 illustrates sunspot structures.

[3] https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/sunspot-regions/usaf_mwl/

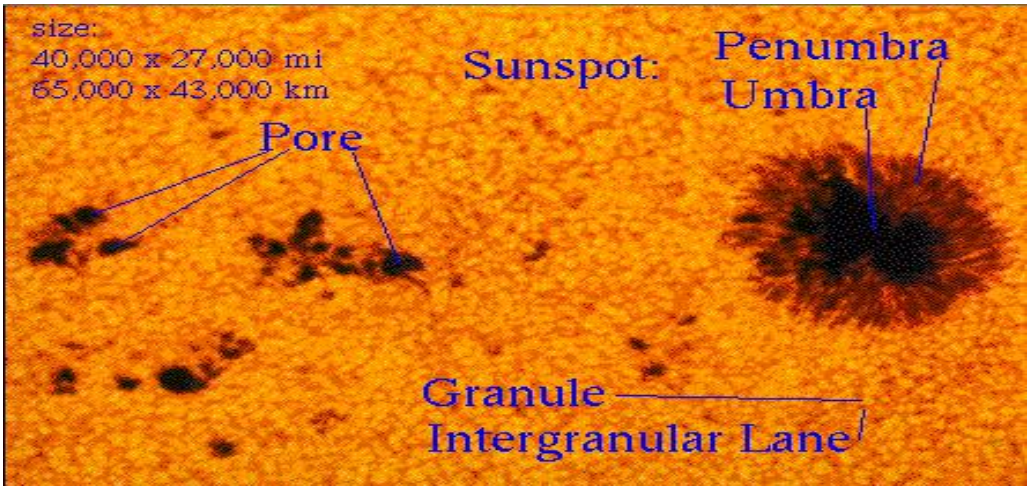


Figure 3: Sunspot structures[4]

1.2.2 Important sunspot features.

As various aspects characterize sunspot groups, it is necessary to identify the common features that can be identified across various types of groups. Important features include the area of every sunspot in the group, and their number and proximity (Solanki, 2003).The presence or absence of a penumbra around the central umbra is another common characteristic (McIntosh, 1990), as is the proximity of spot points to the angular separation between spots, as this indicates whether they will be likely to grow to become a group. Moreover, the 'area' refers to the whole area of the sunspot appearing on the surface of the Sun at the bottom of the photosphere. Finally, the umbra is a darker centre region in the sunspot, and the penumbra is a region surrounding the umbra that is still visually thick, but is a little brighter than the umbra (Foukal 2008). The McIntosh system depends on this combination of elements to classify spots into one of 60 different types (McIntosh, 1990).

1.2.3 Active regions

Older literature describes active regions, or centres of activity, as regions on the Sun that appear in the form of sunspot groups. There are many components, but the Sun's disc and sunspots tend to be the most active regions, and are therefore studied to predict solar activity (T. Colak

[4] https://starchild.gsfc.nasa.gov/docs/StarChild/questions/question17.html

and R. Qahwaji, 2007). The active regions are responsible for a wide range of spectral emissions, such as X-rays and decimetric radio waves. The complex magnetic fields that are linked to active solar regions provide a perfect environment for the creation and release of substantial amounts of energy in the form of solar flares. The launched energy acceleration of atomic particles may cause the Sun's energetic status to change, and this can have a negative impact on systems, such as orbiting satellites and terrestrial systems (Alomari, 2009). The analysis of these data can help in the prediction of solar eruptions that include, for example, solar flares and Coronal Mass Ejections (CMEs), therefore, by enabling predictions, steps could be taken to avoid or ameliorate their effect (T. Colak and R. Qahwaji, 2007).

An active region reaches its maximum age when its ratio of growth stops and the number of spots within the region reaches a maximum number. Meanwhile, flare rates are often reduced when active regions start to decay (Solanki, 2003). However, these processes are linked to the strength of the magnetic field and the gradients in the active region. Active regions decay when the magnetic fields become less complex, which occurs alongside the expansion and fading of plage regions and sunspots (Solanki, 2003).

In this research, active regions are studied from images that are taken by the PICARD satellite at different wavelengths of 215 nm, 393 nm, 535 nm, 607 nm and 782 nm. Additionally, SOHO images are investigated. Continuum intensity images from PICARD/SODISM are mostly used in the analysis of active regions. Recently, different instruments have produced several other types of image, including white-light, H-α (Hydrogen-alpha) and magnetograms. Two types of magnetogram tend to be applied in the analysis of active regions, and these are: longitudinal (line-of-sight), and vector magnetograms. A vector magnetogram is a useful tool for the examination of magnetic fields and flare predictions; it shows the strength and direction of the magnetic field, which are directed along the surface of the Sun (Alomari, 2009). The direction and magnitude of the transverse magnetic field are represented as a vector, while areas where the transverse field is weak or non-existent are displayed as dots (Hu and Sonnerup, 2002).

The polarity of the solar active region depends on whether it is formed in the northern or southern hemisphere. The polarities of the leading and trailing magnetic flux in active regions in the northern and southern hemispheres are the opposite way around (Canfield, 2005). This is a specification of active regions that is rarely disobeyed; therefore, when it is disobeyed, the active region has a reversed polarity configuration, and is often more unsteady and prone to flaring[5]. Figure 4 shows some of the features of the Sun, namely the bright lights and sunspots.

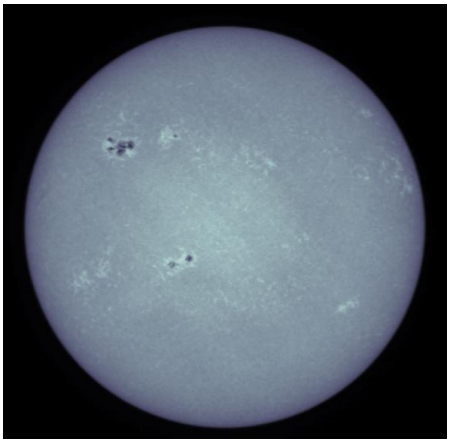


Figure 4: Image by PICARD showing bright regions and sunspots of a wavelength of 393 nm

1.3 Motivation

Due to the importance and potential impact of solar activities, many ground and space-based observatories, such as SDO, SOHO and PICARD, monitor the Sun's activities. These observatories capture different types of data, which are usually analysed manually in a subjective manner by solar experts (Stenning et al., 2013). Therefore, the reported locations and classifications of solar features differ from one observatory to another; the images from the SOHO satellite and those from the PICARD satellite, which provides SODISM images, is one example. The importance of this research lies in the fact that there is currently no reliable, automated system for detecting solar features from SODISM images. To this end, the study aims to develop computer vision algorithms for the segmentation and ultimate extraction of sunspot features from SODISM images. More specifically, a

[5] https://hesperia.gsfc.nasa.gov/sftheory/flare.htm

wavelet transform-based technique is used to solve the underlying problem. Additionally, a filling factor cataloguing technique is proposed to summarise the results, thereby providing the opportunity to compare results in a seamless and easily understandable manner. Furthermore, the proposed technique is rigorously evaluated using comparative studies from different periods.

1.4 Aims and objectives

This thesis makes an important contribution in the realm of the objective interpretation of SODISM data because SODISM images do not have effective automated segmentation methods, or any catalogue summarising sunspot details such as numbers, area, location and their filling factor (FF). This problem leads to searching, and creates a system to remove ambiguity from SODISM data in scientific space research.

The main objective of this research is to provide a model to accurately detect sunspots in SODISM images, and to calculate their corresponding, and to contribute to the automated understanding of the effect of different wavelengths on SODISM images in monitoring a range of features. The outcomes of this research will be mainly the design of new technologies such as a novel enhancement method and a novel SODISM catalogue that will act as a future source for SODISM data details.

This aim can be achieved through the following outcomes:

- To develop a new algorithm to detect solar features in five different SODISM wavelengths and calculate their filling factor.
- To develop a new image de-noising procedure to reduce noise in solar images using a wavelet transform-based method, in order to improve the peak to signal noise ratio (PSNR) of images, and Image quality assessment parameters such as MSE , SSIM , SC, NCC, and NAE are carried out for comparing performance.
- To provide evaluation techniques to assess the accuracy of segmentation.
- To provide visual maps for the segmented solar features.

- To provide comparisons between SODISM images and SOHO images for the same time period.
- To evaluate the proposed techniques using SODISM images in both jpeg and FITS formats.

To achieve these aims, both bright and dark features of sunspots have been detected at different wavelengths. Through its investigation of sunspots, this thesis has added to scientific knowledge of SODISM images, represented, in particular by the fresh methods for enhancing and cataloguing SODISM images. However, the investigation of bright regions has been less successful and, the findings have been discussed to illustrate causes of weakness

1.5 Challenges in working with SODISM Images

The challenges of analysing SODISM images are summarized below:

1.5.1 Data collection

A number of challenges were encountered during the data collection, which included:

From 2014, SODISM images in jpeg formats were not available on the PICARD Website[6] and therefore could not be downloaded; thus, the only valid images were those in Flexible Image Transport System (FITS) format (for wavelength 535 nm). Some images had detrimental features, such as large amounts of noise, striations and ghosts, as illustrated in Figures 5 and 6.

[6] http://picard.busoc.be/sitools/invoquerSva.do?sva=picard&svaAction=searchFiles

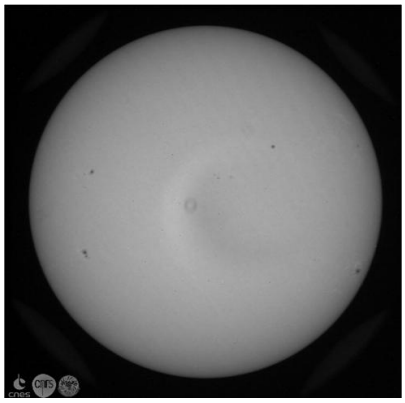


Figure 5: SODISM image at 535 nm on 10th December 2011 at 17:47 showing ghosting effect

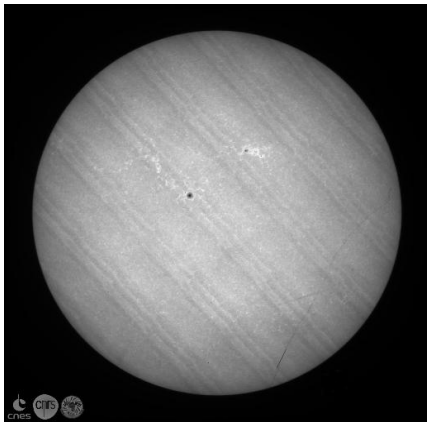


Figure 6: SODISM image at 215 nm on 10th November 2010 at 12:30 showing striations

1.5.2 Research challenges

The main obstacle facing this work was the scarcity of information, published works, references and clarification about SODISM. The principal reason for the lack of information is PICARD's short life and its low-resolution quality images. Initially, the first images that were made publicly available were in a compressed jpeg format, although a second version of the images was later made available in a compression-free Fit format. There was also a limitation in the literature in this area, as many researchers concentrated mainly on the mechanism of the PICARD satellite and its operations (Assus et al. 2008; Etcheto et al. 2011; Meftah et al. 2011; Meftah, et al. 2014; Schmutz et al. 2009). Moreover, papers covering the enhancement or segmentation of SODISM images have not been widely published. Furthermore, the poor quality of the SODISM images made more difficult the design of algorithms to extract the data necessary for sunspot detection. . Additionally, since this is a new research area with a lack of prior studies,

performance comparisons with the findings of other published work have been problematic. Therefore, a cataloguing procedure is proposed in order to help future researchers understand the work, and to facilitate the comparison and evaluation of findings. A further difficulty was experienced due to variations in the data, where the use of different wavelengths led to variations in the image quality. Nevertheless, to the best of the researcher's knowledge, this work is the first scientific research to use SODISM data. Moreover, extensive research is conducted using both versions 1 and 2 of the data releases, which has helped in the collection of hundreds of images, which are therefore relevant data

1.5.3 Original contributions

The original contributions provided in this thesis can be summarised as follows:

- The de-noising of SODISM images was achieved by applying a new enhancement method.
- The extraction of the dark regions from SODISM images helped to show the different wavelength effects during this process. The provision of automated segmentation also helped to detect sunspots at 393 nm, 535 nm, 607 nm and 782 nm.
- The provision of clear map segmentations for sunspots in SODISM images helped to show sunspots on five different wavelengths.
- The study provided a filling factor catalogue for SODISM images, and summarized the whole period of SODISM images in one catalogue.

1.6 Outline of the thesis

This research aims to develop a new image segmentation model for SODISM images and provide new methods to enhance these images using multilevel algorithms. The thesis is organized into six chapters. The first chapter provides the fundamental terminology and basic concepts of the research field, whilst also outlining the aims, objectives, contribution and limitations of the study. Meanwhile, Chapter Two provides an extended

literature review of germane research work on solar sunspots in SODISM images, the automatic techniques used to detect sunspots, and explores wavelet enhancement. Chapter Three specifies the sources of SODISM solar data, and describes the data challenges associated with SODISM, and the different enhancement methods. Furthermore, Chapter Three presents new methods to enhance SODISM images, describes the wavelet transformation methodology, and discusses the practical implementation and evaluation of the segmentation of SODISM features using a new approach that focuses on sunspots. Chapter Four presents the FF calculations of sunspots from SODISM images, and compares them with the results from SOHO images. Chapter Five evaluates the results obtained using the proposed system, and compares the findings with those from other techniques to assess the accuracy of the system. It also provides a comparison between the results for the SOHO satellite images and the PICARD satellite image results. Finally, Chapter Six presents the conclusions of the research and highlights recommendations for future work.

CHAPTER TWO

2. Literature survey

2.1 Image quality enhancement

Image quality enhancement is an important processing step in many computer vision applications, such as medical imaging, astronomy and satellite surveillance (Meftah, et al. 2014). Due to the conditions of the images captured, they may appear blurred and require improvement before visual assessment. This is the case in some of the images downloaded from the Solar Diameter Imager and Surface Mapper (SODISM) hosted on the PICARD Satellite (Bagawade et al. 2012). The aim is to improve the quality of these images (blind corrections), so that the resulting images become more suitable for the extraction of solar features, such as sunspots. To improve the visibility of features in an image, contrast and edge enhancement techniques are used. Alternative techniques, such as bicubic interpolation, are commonly used to increase the number of pixels in an image, which tends to smooth the edges and finer details in the final result (Kumar, Nagaraj and Tech, 2013a). A technology called the discrete wavelet transform (DWT) is used in this enhancement technique to break down the input image into various sub-bands. This is followed by the decomposing of an input low-resolution image into different sub-bands from which high frequency and sub-band images are interpolated. Finally, these images are combined to generate a new image with more visible features. Visual and quantitative comparisons are made between the benchmark, the bicubically enhanced and the DWT enhanced images using both pixel differences and segmented features; for example by examining the sunspot FFs. After it is super-resolved, the main losses in a high-frequency image are the edges. Due to the smoothing caused by interpolation, the DWT works to preserve the high-frequency components of the image (Bagawade et al. 2012).

The simplest interpolation methods used are the nearest neighbour, bilinear and bicubic methods. Each has its own advantages and disadvantages. Nearest neighbour interpolation is the quickest method, but results in significant jagged distortions of the edges. The bilinear method

results in smoother edges, but produces a somewhat blurred appearance. The bicubic method is slowest, but produces smooth edges which also appear as the sharpest (Gonzalez 2002; Mrudula and Lakshmi 2014). Bicubic interpolation is often chosen over its nearest neighbour or over bilinear interpolation when speed is not a problem. Examples in the relevant literature, (Kumar et al. 2013a; Demirel and Anbarjafari 2010) show that it gives better results than the other techniques. It is worth mentioning that these methods were not applied to hyper-spectral SODISM images. Moreover, to improve the quality of the enhanced image and thus preserve details in the image, such as the sharpness of edges, a different enhancement method is essential. In this thesis, a DWT based approach has been employed to preserve the sharpness of edges in the image. Interpolation is a technique that estimates a new pixel value from the neighbouring pixels in the original image (Raju et al. 2014).

In this work, interpolation is based on the use of known data values, which allows the prediction of unknown data values. However, the computational burden increases as the order of the interpolation factor increases (Sangeetha and Krishna 2013); however, this problem may be resolved by interpolating the image in a sub-sampled wavelet domain. The 2-D DWT of an image is performed by implementing first the 1-D DWT along the rows of the image, before down-sampling and then repeating the process along the columns of the results, as shown in the filter band approach illustrated in Figure 7. This operation decomposes the original image into four lower resolution sub-bands referred to as low-low (LL), low-high (LH), high-low (HL), and high-high (HH). The latter three sub-bands occupy the upper-frequency spectrum of the original image, and the resolution enhancement technique is applied at this point in the application (Sangeetha and Krishna 2013).

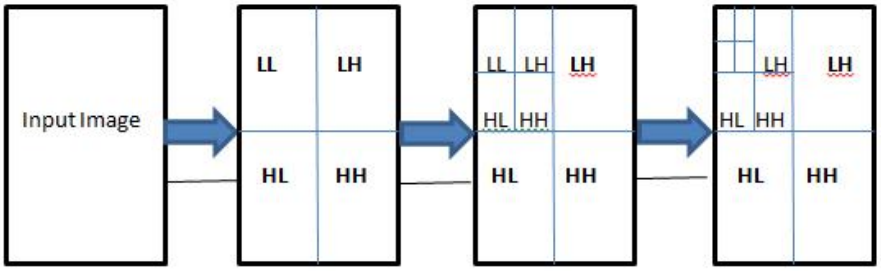


Figure 7: Applying the LL, LH, HL, and HH sub-band DWT three times on the lowest sub-band starting from the input image

One transformation related to the DWT, which is used in the present application as well as in many other image processing applications, is the stationary wavelet transform (SWT) (Narayana and Nirmala 2012; Li and Orchard 2001). Here, the low-frequency and high-frequency sub-bands of the SWT are not down-sampled, and have the same size as the input image.

- Image enhancement aims to enhance the visual quality of an image to improve the processing of regions of interest. Techniques, such as noise reduction, contrast stretching, smoothing and edge enhancement, are usually carried out during the enhancement.

- The condition of solar images is intrinsically different to other types of images as they contain unwanted noise that occurs both during image acquisition, and in the image transfer process to Earth-based stations.

2.2 Different sources of noise in images

One of the most prominent problems to hinder the accuracy of computer vision algorithms is image noise, and this is particularly significant in satellite images. The presence of noise in satellite data introduces irrelevant and undesired information that can lead to the false identification of important features. Thus, noise is defined here as any corruption of the image signal occurring due to external disturbance. Furthermore, variations in image intensity represent random noise, which is visible as pixel fluctuations within an image. Some of this noise may appear as a result of the basic, physics-like photon nature of bright or thermal energy within the image sensors. The main sources of noise in digital images include the following (Verma and Ali 2013):

1. Environmental conditions.

2. Insufficient light levels and sensor temperature.

3. Dust particles.

4. Interference in the transmission channel.

2.3 Noise detection

As previously mentioned, noise is defined as any corruption of the image signal, and occurs due to external disturbance (Verma, 2013). When images are transmitted electronically from one place to another, for instance via satellite or any other transmission device, signal synchronization problems and magnetic interference can result in disturbances that eventually lead to image noise. The image noise appears as output in different forms depending on the kind of signal disturbance (Verma, 2013).

2.3.1 Noise identification

Due to the adverse effect of noise, it is important to identify its source (Boyat and Joshi, 2015). While the signal is often easy to compress, noise, is not so readily compressible (Pambrun and Noumeir, 2015). Although noise identification is not straightforward, in many instances, the instrument employed to acquire the image provides the necessary information that can be used to detect and identify the properties of underlying noise (Sciacchitano, Dong and Hansen, 2017). In most cases, the effect of noise is estimated on a pixel-by-pixel basis. Although noise distribution varies, most researchers primarily assume that digital image noise (such as that caused by temperature or transmission impairment, and variation illumination) is normally distributed; this is popularly known as Gaussian noise (Bosco et al., 2003).

It is worth mentioning that, for normally distributed noise, its second statistical moment (or the standard deviation) plays an important role in its identification. Several methods have been used to estimate the standard deviation of Gaussian noise in digital images, and the simplest, and probably best, is the averaging method (Kumar and Gupta, 2012). This consists of filtering the data with an average filter (known also as arithmetic mean filter) and then subtracting the result from the original image, which gives the

means to measure the noise associated with each pixel (Kumar and Gupta, 2012).

2.3.2 Noise thresholding parameters

Digital image noise can be eliminated using a variety of techniques, including (but not limited to) the application of a smoothing filter, which may either be linear or nonlinear (Verma, 2013). However, reducing noise in images might need a 'brute-force' technique, called noise thresholding (Mishra and Mitra, 2014). This involves setting pixel values to zero (or to a convenient lower value) for all pixels under a certain limit that is given in terms of a previously measured noise level. This provides the assumed Gaussian distribution of noise over the whole image, and the thresholding noise level can be derived from sigma (standard deviation) values via iterative calculations (Chan and Zhou, 2003).

2.4 Noise models

The equation for a noisy image, as stated by Lysaker et al. (2003), can be expressed as:

$$C(x+y) = A(x+y) + B(x,y) \quad (1)$$

where *A(x ,y)* represents the pixel value in the original image and *B(x ,y)* is the noise in the image while *C(x ,y)* is the resulting noise image. Figure 8 shows typical noise model types.

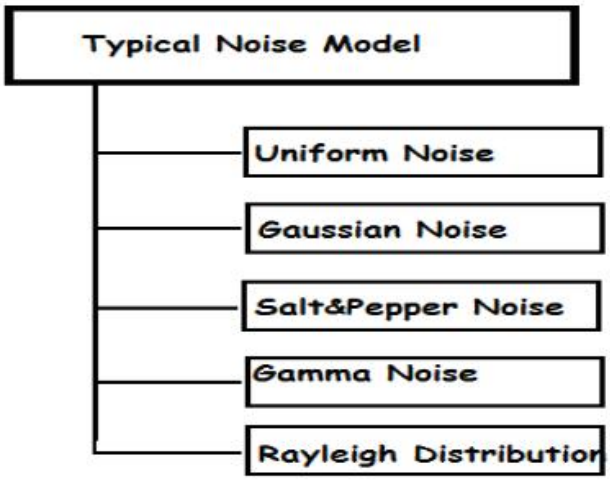


Figure 8: Typical noise model type (Kamboj and Rani, 2013)

2.4.1 Types of noise

1. ***Uniform noise:***

The quantization of pixel image to a number of distinct levels is a key reason for uniform noise. Usually it has a uniform distribution where the level of grey values for noise is uniformly distributed across a specified range. However, this noise is used to estimate image restoration algorithms because it provides unbiased, or the most neutral, noise (Singh, 2016). An example of uniform noise is shown in figure 9.

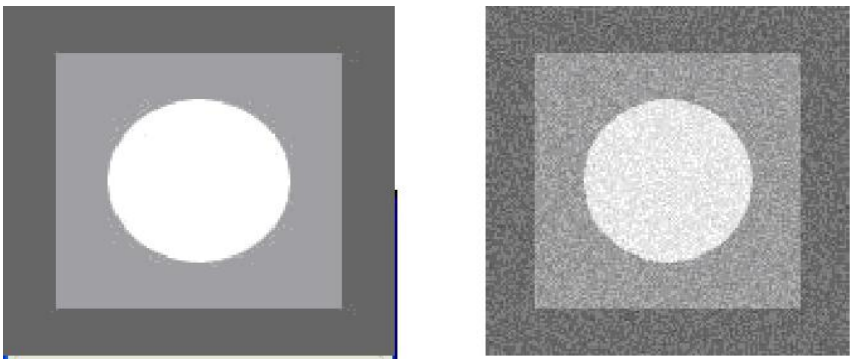


Figure 9: Example of uniform noise (Verma and Ali, 2013)

2. ***Gaussian or amplifier noise:***

Normal noise, or Gaussian noise, is additive in nature following a Gaussian distribution. It is a probability density function [PDF] of the normal distribution, and is the sum of the true pixel value alongside a value of random Gaussian distributed noise that represents each pixel in the noisy image. The intensity of the pixel value at each point determines the noise. The PDF of the Gaussian random variable is given as follows (Kamboj and Rani, 2013):

$$P(x) = 1/(\sigma\sqrt{2\pi})e^{(x-\mu)2}/2\sigma^{2} \; -\infty<0<\infty \quad (2)$$

where *P(x)* is the Gaussian distribution of noise in the image; and *μ* and *σ* are the mean and standard deviations respectively. Figures 10 and 11 respectively, show a Gaussian histogram and an example of the effect on an image.

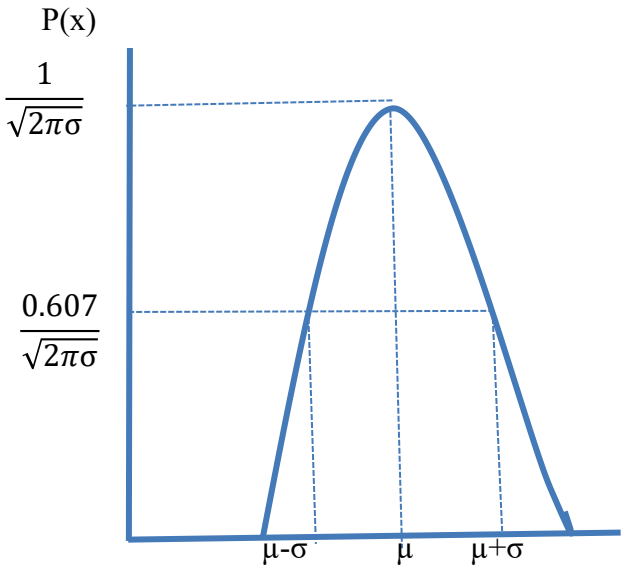


Figure 10: PDF of Gaussian Noise (Kamboj and Rani, 2013)

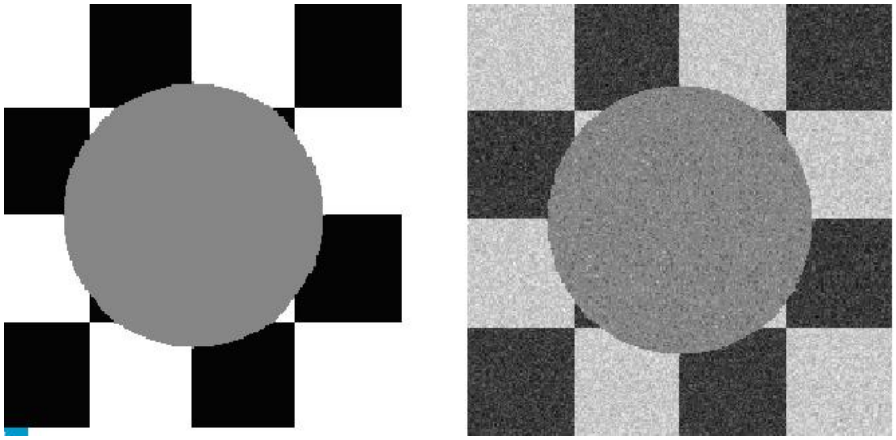


Figure 11: Example of Gaussian noise

3. ***Salt-and-pepper noise:***

Shot noise, data drop also known as salt-and-pepper noise, (other names include impulse noise or spike noise) may be caused by factors such as timing errors during the process of digitization (Singh, 2016). This can result from faulty memory positions or malfunctioning pixel elements in the camera's sensors. In salt-and-pepper noise, two possible values exist, which are *a* and *b,* and the probability of each is less than 0.2. If the numbers are greater than 0.2, then the noise will swamp the image. For an 8-bit image, the typical intensity value for salt-noise is 255, and pepper noise is 0 (Leavline, Asir and Singh, 2013).

There are three main reasons for salt-and-pepper noise (Gabhel and Hiradhar, 2014):

- Memory cell failure.
- The malfunctioning of the camera's sensor cells.
- Synchronization errors during transmission or in the image digitization.

Impulse noise:

$$p(z) = \begin{cases} p_a & for \quad z = a \\ p_b & for \quad z = b \\ 0 & otherwise \end{cases}$$

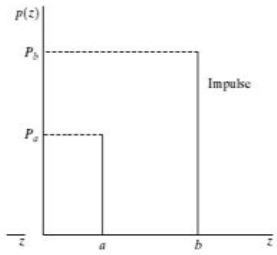


Figure 12: PDF of impulse noise (Valens, 1999)

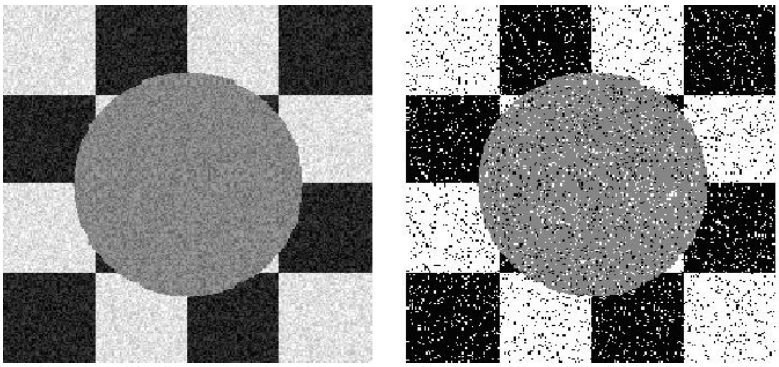


Figure 13: Example of impulse noise

4. Poisson noise (photon noise)

When the sensor identifies insufficient numbers of photons to provide detectable statistical information, the outcome is Poisson, or shot photon noise. This has a root mean square value proportional to the square root intensity of the image. Usually, independent noise values affect different pixels. However, the signal can be corrupted in different proportions, particularly by photon, and other sensor-based noise (Singh, 2016).

5. Speckle noise

Random value multiplications with pixel values cause this type of noise. It is also called a multiplicative noise, and it can be expressed as:

$$J = I + n*I \quad (3)$$

where J is the speckle noise distribution image, I is the input image, and n is the uniform noise image with mean o and variance v. This form of noise

tends to cause deteriorations in the quality of active radar and synthetic aperture radar (SAR) images. Speckle noise is usually observed in radar sensing systems and images, when the signal from the object is reflected with a size less than, or equal to, a single image processing unit; moreover, it also shows sudden fluctuations (Singh, 2016).

2.3 Image de-noising

The initial step, before the application of image processing tools, is to remove the noise (unwanted effects) from it (de-noising). In fact, the greatest challenge is designing an efficient methodology to remove noise from images. Noise removal is a major problem in computer vision and image processing, and elimination is difficult because it can cause image blurring. A suitable filter must be applied to remove the noise from an image; however, identifying the kind of noise present in a degraded noisy image is a priority. Therefore, it is necessary to apply an appropriate filter and correct any degradation in the image. Various de-noising algorithms are available, and the best can remove unwanted noise completely from the image while preserving the details and avoiding distortion.

Depending on the kind of noise present in an image, the types of noise removal algorithms can differ. Once the types of noise are identified, various algorithms can be applied to erase it, or certain kinds of noise can be removed. There are two types of de-noising methods: the first has the advantage of rapid operation but does not fully preserve the details of the image. The second type involves non-linear methods, which retain details of the image at the expense of speed in completion.

2.4 Filtering techniques

There are two main filtering techniques, which are as follows:

- ***Linear filters:***

  As previously stated, linear filters have been used to erase certain types of noise. Gaussian or averaging filters are suitable to achieve this goal. However, some filters tend to blur sharp edges, damage pixels and other fine details, or perform inefficiently in the presence of signal-dependent noise.

- ***Non-linear filters:***

  There are many non-linear median kinds of filters, such as rank conditioned, relaxed median, weighted median and rank selection methods. These have been developed to eliminate the shortcomings of the linear filter (Hardie and Barner 1994; Khan 2016; Shmulevich and Arce 2001) (Singh, 2016).

2.4.1 Different types of linear and non-linear filter

1. Mean filter:

The mean filter is a sliding-window that replaces the central value in the window (Singh, 2016). It replaces it with the mean of all of the pixel values in the kernel or window (Singh, 2016). The window is usually square but can be any shape. A simple spatial filter is considered in Figure 14, where the central value, which was previously 1 in the unfiltered value, is replaced by the mean of all nine values, namely 5. Mean filters are best for Gaussian noise and uniform noise, and the advantages are that they can be used to remove impulse noise and are easy to implement. The disadvantage is that the details in the image may become faded, which means that the filter does not preserve all of the detail.

Figure 14: Example of the mean filtering of a 3x3 kernel of values (Kamboj and Rani, 2013)

2. Median filter:

The median filter is considered to be a powerful and simple non-linear filter, which is based on ordering statistics. It is easy to implement with smooth images (Kamboj and Rani, 2013). The median filter is used to reduce the amount of variation in intensity among pixels (Singh, 2016). In this filter, the central pixel is replaced by the median value (Kamboj and

Rani, 2013). Calculations in the median filter start by sorting all of the pixel values into ascending order and replacing the pixels with the middle pixel value. Therefore, this method uses the average of the neighbouring pixel values. If the neighbouring pixel value contains an even number of pixels, then the average of the two middle pixel values is chosen as a replacement. The median filter produces the optimum result when the percentage of impulse noise is less than 0.1 %. However, increasing impulse noise levels have negative impacts on the effectiveness of the median filter. The median filter is considered easy to implement and is used for de-noising many types of noise from an image. On the other hand, a median filter can lose image details, such as corners and lines. Moreover, the performance of this filter is unacceptable for some cases, such as when removing signal-dependent noise. However, different variations of median filters have been developed to overcome these weaknesses and provide better results (Kamboj and Rani, 2013).

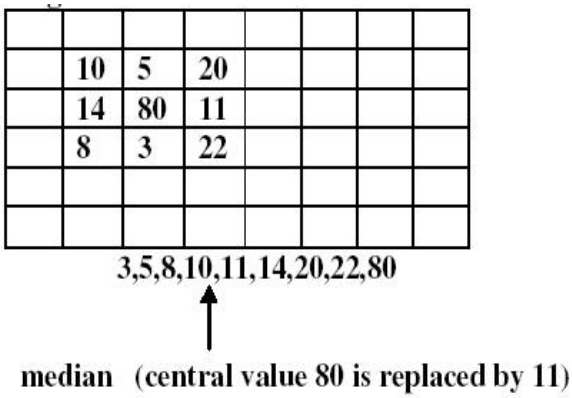


Figure 15: Example of median filtering of a 3x3 kernel of values (Kamboj and Rani, 2013)

3. Adaptive filter

This filter does not always exhibit the same behaviour, and the results radically change depending on the statistical characteristics of the image region encompassed by the filter region (Singh, 2016). The BM3D is one kind of adaptive filter that uses a non-local image modelling technique, depending on the adaptation. The algorithm for de-noising in this filter has the following three steps:

- **Analysis**: the initial step summarizes to make similar image blocks, and then stacks these blocks together to form 3-D data arrays. These are then de-correlated using an invertible 3-D transform.
- **Processing**: this step uses hard thresholding to filter the previous 3-D group spectra.
- **Synthesis**: this step depends on the inverse of the filtered spectra and gives estimates for each block in the group, returning block-wise estimates to their original positions. The final image is then obtained by calculating a weighted average of all the block-wise estimates.

4. Wiener filter:

The purpose of the Wiener filter is to filter out noise that has corrupted a signal. The basis of this filter is a statistical approach. Unlike other filters, which are designed for a desired frequency response, this filter reduces the mean square error (MSE) as much as possible. The Wiener filter also has the capacity to decrease the degrading function and noise. Moreover, this filter adopts the linear estimation of the original image, which depends on a stochastic framework, and a linear time-invariant filter that gives a similar original image (Singh, 2016).

2.4.2 Performance parameters:

2.4.2.1 Mean square error (MSE):

The mean square error (MSE) can be used to compare original images with compressed images, and the lower MSE value gives a better quality compressed image. The MSE is known as the cumulative square error between the encoder and the original image, which can be calculated as follows (Khobragade et al. 2014).

$$MSE = \frac{1}{mn} \sum_{0}^{m-1} \sum_{0}^{n-1} \|f(i,j) - g(i,j)| \quad (4)$$

where *f* is the original image and *g* is the compressed image. The dimensions of the images are m *x n*. Thus the value of *MSE* should be as low as possible for effective compression.

2.4.2.2 Image structural similarity based metrics (SSIM):

The mean square error (MSE) and the peak to signal noise ratio (PSNR) are widely used for full-reference image quality measurement, but SSIM is a more accurate means of measuring quality. (Varnan et al., 2011) SSIM measures the quality of the black and white image but has a disadvantage in the case of a highly blurred image, and in fact fails in this case (Varnan et al., 2011).

SSIM is given by equation number 5.

$$SSIM = \frac{(2 \times \bar{x} \times \bar{y} + C1)(2 \times \sigma_{xy} + C2)}{(\sigma_x^2 + \sigma_y^2 + C2) \times ((\bar{x})^2 + (\bar{y})^2 + C1)} \quad (5)$$

where *C1* and *C2* are constants. $\bar{x}$, $\bar{y}$, σ_x ,σ_y and σ_{xy} are given by:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (5.1)$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \quad (5.2)$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})^2 \quad (5.3)$$

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N} (y_i - \bar{y})^2 \quad (5.4)$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}) \quad (5.5)$$

2.4.2.3 Peak signal to noise ratio (PSNR):

The PSNR is the ratio of the maximum possible power of a signal to the power of the distorting noise that affects the quality of representation:

$$PSNR = 20 \log_{10} \left(\frac{MAX_f}{\sqrt{MSE}} \right) \quad (6)$$

where MAX_f is the maximum signal value that exists in the original “known to be good” image. The ‘bits per pixel’ (BPP) is defined as the number of bits required to compress each pixel; this should be low in order to reduce storage requirements (Khidse, 2014).

2.4.2.4 Signal-to-noise ratio

The SNR is defined as the power ratio between a signal (meaningful information) and the background noise (unwanted signal):

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}}, \quad (7)$$

where *P* is the average power. Both noise and power must be measured at the same points in the system, and within systems with the same bandwidth.

An alternative definition of SNR is the ratio of the mean to the standard deviation of a signal or measurement:

$$\mathrm{SNR} = \frac{\mu}{\sigma} \quad (8)$$

where μ is the signal mean or expected value, and σ is the standard deviation of the noise, or an estimate thereof. It is a spare definition, and only active for variables that are always non-negative. It tends to be used in image processing where the SNR of an image is calculated as the ratio of the mean pixel value to the standard deviation of the pixel values over a given neighbourhood.

2.4.2.5 Normalized absolute error (NAE):

Normalized Absolute Error (NAE) represents by dividing the absolute difference between the original, and the enhanced image by the original image. The high value of NAE means the image is poor (Khidse, 2014).

$$\boldsymbol{MAE} = \frac{\sum_{i=1}^{m}\sum_{k=1}^{n}|X_{i,k}-X'_{i,k}|}{\sqrt[2]{\sum_{i=1}^{m}\sum_{k=1}^{n}X_{i,k}}} \quad \mathbf{(9)}$$

where *M* and *N* is a dimension of images and *i* and *k* are indexing of two-dimensional image arrays, the *x* and *x'* represent the original image and enhanced image.

2.4.2.6 Normalized cross-correlation (NCC):

Normalized cross-correlation (NCC) is used for image-processing applications in which the brightness of the image and template can vary due to lighting and exposure conditions. Normalizing images can be done in

every step by subtracting the mean and dividing by the standard deviation. The NCC measures the similarity between two images as a function of the lag of one relative to the other, which is proven to obtain more accurate matches (Khidse, 2014).

$$NCC = \frac{\sum_{i=1}^{m}\sum_{k=1}^{n} X_{i,k} - X'_{i,k}}{\sqrt[2]{\sum_{i=1}^{m}\sum_{k=1}^{n} X_{i,k} . X'_{i,k}}} \quad (10)$$

where *M*, *N*,*I*, *k*, *x* and *x'* represent the same as in equation 9.

2.4.2.7 Structural content (SC):

The Structural content is the ratio between the square of pixel values of the original image and the enhanced image (Khidse, 2014).

$$SC = \frac{\sum_{i=1}^{m}\sum_{k=1}^{n} X_{i,k}^2}{\sqrt[2]{\sum_{i=1}^{m}\sum_{k=1}^{n} (X'_{i,k})^2}} \quad (11)$$

where *M*, *N, I*, *k*, *x* and *x'* represent the same as in equation 9

2.5 Wavelet transforms

The wavelet transform is similar to the Fourier transform, which decomposes the signal into sines and cosines, while the wavelet transform uses functions that are centralized in both real and Fourier space. This is why the term 'wavelet transform' is used in very different situations and applications. Mathematically, wavelet transforms (WTs) perform signal analyses when a signal frequency varies over time. The WT is capable of providing time and frequency information simultaneously, and hence a time-frequency scale for the signal is represented. The wavelet transform, or wavelet analysis, is developed to overcome the shortcomings of the Fourier transform.

For certain types of signals and images, wavelet analysis provides more precise information about the signal data than other analysis techniques. In wavelet analysis, the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and a spectrum is calculated for every position. The process is then repeated many times with a slightly shorter (or longer) window for every new cycle. The final result is a collection of time-frequency representations for the signal,

all with different resolutions. Due to this collection of representations, a multi-resolution analysis can be applied. In the case of wavelets, the time-frequency representations are not seen, but the time-scale representations are taken into consideration, with scale being (in a sense) the opposite of frequency. This is because the term frequency is reserved for the Fourier transform (Lakshmi, 2013). The wavelet transform estimates sub-bands based on a filter bank. The outputs of the different filter stages are the wavelet and scaling function transform coefficients. An approach using the wavelet transform is implemented with a version of the 'so-called' *à trous* algorithm.

2.5.1 Á trous wavelet transform and multiresolution support:

The á trous wavelet transform algorithm is well-known as a non-orthonormal multi-resolution algorithm, for which the discrete wavelet transform concerns 'algorithme à trous' (González-Audícana et al., 2005). This is from the French, where 'trous' means holes, and refers to the insertion of zeros in the filters. An à trous algorithm can represent a distinct approach compared to the continuous wavelet transform, (CWT) and it is a powerful tool for the multi-scale analysis of images. Furthermore, enhancement tasks and noise reduction can be performed with this algorithm, and it modifies the decomposition scheme of the standard discrete wavelet transform (DWT) by modifying the low-pass and high-pass filters at each repeated level. Once the WT has provided band-pass bands with bandwidth, and the scaling function provides the low-pass band, it can be concluded that a wavelet transform is the same as a sub-band coding plane. Using a constant-Q filter bank in this kind of analysis is called multi-resolution analysis. The multi-resolution support of an image is described in a logical or Boolean way, whereby if $m(j,x,y)=1$ then an image contains information at scale j at the position (x,y).

The à trous WT performs a hierarchical decomposition of an image into a series of scale layers arranged according to rank. Each layer contains only structures within a given range of characteristic dimensional scales in

the space of a scaling function. The decomposition is performed through a number of detail layers defined at increasing characteristic scales, plus a final residual layer that contains the rest of the structures. By isolating the desired image structures within the specific detail layers, detail enhancement can be carried out with high accuracy. Likewise, if noise occurs at some specific dimensional scales in an image, it can be reduced or removed without affecting the significant structures by isolating it into appropriate detail layers.

2.5.2 Image resolution enhancement using discrete and stationary wavelet decomposition

The discrete wavelet transform (DWT) is applied to input images depending on the satellite image resolution enhancement technique used by the interpolation of the high-frequency sub-bands. The applied resolution enhancement technique uses DWT to decompose the input image into different sub-bands (Harikrishna and Maheshwari 2012). The resolution-enhancement technique applies interpolated DWT high-frequency sub-band images and an input low-resolution image. From this, a new resolution-enhanced image is generated by combining the previous images through the application of the inverse DWT (IDWT) (Khaire and Shelkikar 2013).

These techniques have been applied to many remote sensing and satellite images to enhance the image resolution. (Harikrishna and Maheshwari, 2012) showed the effectiveness of this method compared to other methods as visual points of view were included. The paper shows the results of satellite images for the proposed method and a comparison with a bicubic interpolation. Moreover, the images are enhanced through the use of WZP image resolution enhancement. The algorithm applied is shown in Figure 16.

The peak signal-to-noise ratio (PSNR) was applied in order to achieve some quantitative results for comparison. The PSNR can be determined using the following formula:

$$PSNR = 10\log\left(\frac{R^2}{MSE}\right) \quad (8)$$

where R represents the maximum fluctuation in the input image (255 in this study), and the MSE represents the between the given input image and the original image.

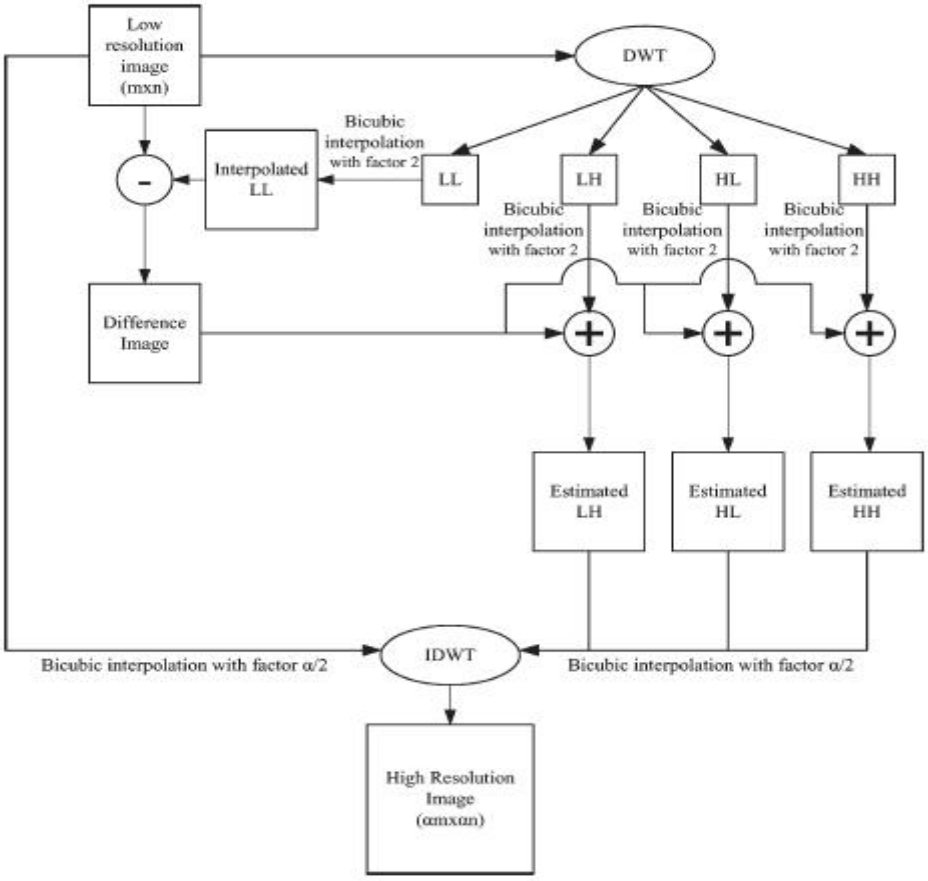


Figure 16: Block diagram of resolution enhancement algorithm (Demirel and Anbarjafari, 2011a)

2.6 Segmentation

2.6.1 Background

Solar images contain many types of data collected from different satellites, including PICARD. This data provides opportunities to predict the future effects of the Sun on Earth, which is important, as powerful solar activities can have damaging effects on a wide range of technologies, including satellites, airlines, computer memory, telecommunications, GPS signals, astronaut safety, power distribution systems, and the high-frequency communication used by satellites. This increasing dependence on

technology means that humans are becoming more vulnerable to the potential damage that could be caused by solar activities. A lot of noise information - distortion and pixel noise - is present in many of the PICARD astronomical solar images. However, if the unwanted noise can be reliably estimated by an image restoration, then useful information can be extracted to assess the features of the images. Furthermore, turbulence in the Earth's atmosphere will affect Earth-based solar observations in different ways. In practice, to study the fine structure of the solar surface, it is first necessary to implement image enhancement processes, and thereby improve the visibility of many categories of solar features, such as sunspots (Adipranata et al. 2103 ; Christopoulou 2002).

Dark areas at a lower temperature than their surroundings are called sunspots. The temperature of sunspots is lower than the average temperature of the solar surface (Curto, Blanca and Martínez, 2008). A sunspot is composed of an umbra and penumbra. It is a prerequisite for the understanding of spatial atmosphere and solar physics to segment and extract sunspot features using automated detection techniques. The detection process includes two major steps, namely segmenting and extracting sunspots (Yu, Deng and Feng, 2014).

Approximately seven categories of method can be applied to detect solar features (Zharkova et al., 2005) and these are listed below:

- Thresholding approaches depend on image histograms, and are good for simple images that have features superimposed on a constant background intensity.

- Edge-based methods are easy to apply but provide incomplete edges, meaning that another process to link the solar image is needed. This method uses an intensity function spatio-temporal with first or second derivatives.

- Region-growing methods are efficient for homogeneous regions if the assumptions are valid. The initial starting pixel and a criterion for merging adjacent pixels are required.

- Hybrid region-growing and edge-based techniques. These methods are based on the combination of two or more approaches where the choice depends on the advantages and properties of each approach in achieving the best segmentation results. The disadvantage of this method is its complex processing and the consequent time-consuming process involved.

- Artificial neural network methods require specialist training to enable the recognition of features that do not explicitly determine the criteria.

- Global information methods use Bayesian inference and Hough transforms in the segmentation process. This method is complex and takes a long time to run but provides good results.

- Miscellaneous methods include data clustering, simulated annealing, and data mining, as well as tree/graph-based methods, model-based segmentation, graph partitioning, watershed transformations and multi-scale segmentation.

The previous approaches are the most popular methods for finding object regions in grey-level images, and could also be suitable for colour segmentation (Demirel and Anbarjafari, 2011b).

Image segmentation is the process of dividing a digital image into several segments as sets of pixels. The goal of segmentation is for the regions to show meaningful parts of the image which makes analysis easier. In other words, image segmentation is the breakup of an image into categories that correspond to various objects or parts of objects. Every single pixel in an image is a pointer to one of a number of categories.

The aim of image segmentation is to detect the object's location and its boundaries, which may include lines or curves in the image. In other words, the segmentation of an image is used to determine and assign a label to each pixel in the image, and all pixels with the same label will have similar characteristics (Pratt, 2000). Segmentation, therefore, has two objectives (Srinivasan, 2008). The first objective concerns the further analysis, where the image is decomposed into parts. In some cases this is considered to be

simple because the environment can be controlled, and so the segmentation process will be reliable in only extracting the target parts. An effective example of this is the segmentation of a human face from a colour video image. However, in complex cases, difficult challenges will require extra processing to achieve goals. The second objective is to achieve a change of representation. In this case, further organization of the pixels in the image will be required to ensure higher-level units are more flexible for the next process, which includes analysis. The biggest challenge is for the segmentation to achieve and perform various different domains by applying general bottom-up approaches, without any specific domain-based knowledge (Srinivasan, 2008).

Many general multi-purpose techniques and algorithms have been developed for meaningful segmentation. Thus, there are no rules or general solutions to image segmentation problems, and these techniques often work together as domain knowledge in order to identify an effective solution for an image segmentation problem. In other words, image segmentation needs to be addressed from a wide variety of perspectives.

The shape and layout features depend on the accuracy of the segmentation, which is why it is considered to be the most important process in image retrieval. In the last few years, research into image segmentation has developed, but there are no general algorithms for segmentation techniques that can work for all images (Srinivasan, 2008). Lybanon et al. (1994) proposed a morphological operation that used an opening and closing approach to image segmentation. These methods are applied to different types of images, such as optical astronomical images, infrared ocean images and magnetograms. However, whilst these are successful for scientific image types, they are ineffective for more complex, natural-scene images. In comparison, Hansen and Higgins (1994) used watershed analysis and relaxation labelling, where the former subdivided an image into catchment basins whose classifications were then refined and updated via relaxation. A key characteristic of relaxation labelling is its robustness to filtering noise. In comparison, Li et al. (1994) proposed fuzzy entropy based segmentation approach, based on the fact that the local entropy maximum corresponds to

the uncertainties among various regions in the image. This method works effectively for images whose histograms do not have clear peaks and troughs. Although these methodologies are automatic, facilitating the time-efficient use of algorithms to extract boundaries from a large number of images, human assistance is still required for an unconstrained domain, or for non-preconditioned images because automatic segmentation is not completely reliable.

In comparison, Samadani and Han (1993) proposed a computer-assisted boundary extraction method, which mixes manual human inputs with the image edges obtained by computer. A new approach by Daneels et al. (1993) improves the active contours method which is dependent on the user's input, where the algorithm first uses a greedy procedure to submit a fast initial convergence.

Applying a segmentation process for complicated structure images is considered the most difficult problem in image processing, and has been an interesting active area of research for several years (Zharkova et al., 2002). Various techniques have been devised, and the most significant difference in their method is in their classification, which is somewhat subjective. The most relevant solar processing approaches concentrate on techniques associated with feature recognition within order image segmentation. These depend on non-overlapping regions, and are based on intensity, multispectral ratio, texture, motion, and so forth. However, segmentation, in general, is divided into five main techniques, which are histogram-based, region-based, edge-based, artificial neural networks, and explicit-model based segmentation. Furthermore, each of these techniques have sub-techniques (Zharkova et al., 2002).

Ashamari et al. (2015) provide automated technologies to identify solar activity features from large databases. They achieve this by presenting a series of enhanced segmentation algorithms which detect and calculate the area coverages of specific magnetic features from MDI intensitygrams and magnetograms (Ashamari, Qahwaji, Ipson, Scholl, et al., 2015). Moreover, they provide binary segmentation maps along with the solar area coverage of the features. Their method depends on the Automated

Solar Activity Prediction (ASAP) program (Colak and Qahwaji, 2008) for which they develop new segmentation algorithms that are integrated into ASAP to detect and identify faculae and networks. Thus, they use a Neural Network (NN) to group sunspots and discover active regions based on sunspot properties and magnetic regions (Ashamari et al., 2015). However, despite the fact that many segmentation methods might be applied to SODISM images, it is still a challenge to achieve acceptable results in this field. Nevertheless, the analysis of SODISM images and creation of segmentation maps could lead to the classification of active regions and the production of good space weather prediction systems.

2.6.2 Features of the solar images

As described previously, the decision making process for every image pixel as to whether it is part of an image feature, is called feature detection, or segmentation. The results of feature detection or segmentation methods are subsets of the image, and usually appear in the form of isolated points. Once the solar disk has been detected, the next step is limb darkening removal, which involves the application of suitable intensity filtering. Although limb darkening removal is important, it is not a problem to detect active regions because limb darkening darkens the solar background. However, most solar images have limb-darkening problems, which involve a gradual reduction in the brightness of the disk of the Sun, as noted from its centre to its border edge. Limb darkening occurs because the solar atmosphere increases in temperature with depth (Qahwaji and Colak, 2005a). The algorithm provided by Qahwaji (2005) helps to determine the exact positions of active regions, and filaments can be useful to detect other features, such as sunspots if the intensity filtering is modified (Qahwaji and Colak, 2005a). Filament detection algorithms aim to determine the presence of a filament in solar images, and thus detect its characteristics; this includes properties, such as the centre, area, length, head start-point, tail end-point and filament border points. This process helps in other operations, such as merging broken filaments in order to restore the actual size of the detected filaments (Al-Omari et al. 2010).

Solar images can present particular difficulties for image analysis, some of which are inherent, such as white-light solar images that suffer a limb darkening effect (Hubrecht 1913, Heinz and Dietrich 1994). The intrinsic quality of the image affects sunspot detection, due, for example, to the quality of the optical lens system and the equipment used which determines the image quality and the appearance of details. The light intensity in sunspot images also depends on their location on the solar disc. Additionally, sunspot shapes do not maintain the same pattern throughout their lives, which increases the challenges in working with such data. Moreover, the difficulties increase with exposure times and the sensitivity of the sensor used for the Sun's luminosity.

Many factors affect the images of the Sun taken from the Earth's surface, such as the changing distance between the Sun and the Earth, the angle of solar rays captured in images, the veiling effect of clouds, and the transparency of the Earth's atmosphere. However, the data used in this thesis were collected by the PICARD satellite, which only has a minor level of dependence on the distance between the satellite and the Sun (in space). In practice, the probability of extracting sunspot features depends on the accuracy of the captured images. Data collected from a PICARD satellite present many challenges because the images are not always perfect. Small sunspots and small details might disappear for several reasons, such as unfocused images, enlarged sunspots, and reduced intensity and contrast, which can increase the signal-to-noise ratio (Curto, Blanca and Martínez, 2008). Furthermore chromatic and spherical aberrations in the optical lens may also produce distorted images with similar effects to those of unfocused images (Curto, Blanca and Martínez, 2008). Furthermore, shadows are sometimes also interpreted as sunspots, due to the build-up of water and dust on the optical elements, or on the charge-coupled device (CCD) sensors (Curto, Blanca and Martínez, 2008). Moreover, corrupted CCD cells may introduce errors to the intensity values of their pixels, and thus a complex filtering process may be needed to clear the background images and retain the information related to sunspots at the same time.

Massive amounts of data are available on the features observed in solar activity, such as flares, active regions, filaments, sunspots, magnetic neutral lines, and coronal mass ejections. These can be analysed using automatic tools (Zharkov et al. 2009). Several methods have been proposed using Bayesian image segmentation (Turmon, Pap and Mukhtar, 2002), mean-field fast annealing (Bratsolis and Sigelle, 1998) and even neural networks (Fernandez et al. 2002).

Interpolation methods can be used to enhance image resolution, which include the nearest neighbour, bilinear, and bicubic approaches, each of which has its own disadvantages. The nearest neighbor approach is computationally efficient but not as accurate as bilinear or bicubic interpolations. A bilinear interpolation results in smoother edges, but with a somewhat blurred appearance overall. In comparison, a bicubic interpolation is more sophisticated than the other two techniques and produces smoother edges and less blurring, but is the slowest method (Gonzalez 2002, Bagawade P. and Bhagawat S. 2012). Thus, bicubic interpolation is the best choice compared to bilinear or nearest neighbour interpolation, when time is not an issue. From examples in the literature, bicubic interpolation gives better results than the other techniques (Mrudula and Lakshmi 2014, Kumar et al. 2013a). Studies that describe various schemes available for resolution enhancement are reviewed below.

Demirel and Anbarjafari (2010) proposed a technique which uses DT-CWT to decompose satellite images into different sub-band images. The benchmark image is interpolated with half of the factor used for the interpolation of the high-frequency sub-band images after which IDT-CWT is used to generate super-resolved combined images. Piao et al. (2007) used inter-sub-band correlation in the wavelet domain to enhance image resolutions, and this method utilized the correlation of sub-bands with different sampling phases in the DWT.

Li and Orchard (2001) developed a novel edge-directed interpolation algorithm which provides significant improvements in the visual quality of interpolated images compared to linear interpolation. In comparison, Atkins et al. (2001) created a new method for optimal image scaling called

resolution synthesis (RS). This provides a simple derivation where the RS generates the minimum mean-squared error (MMSE) estimate of a high-resolution image, given the low-resolution image. Temizel and Vlachos (2005) posed an approach which estimates the local edge orientation from a wavelet decomposition of the low resolution (LR) image, and this information is used to control cycle spanning (CS) parameters.

2.7 PICARD satellite and previous related works

2.7.1 Work related to SODISM

The problem of sunspot identification using SODISM is relatively new; from 2010 to 2014, the PICARD satellite recorded over one million pictures of the Sun and a horde of other measurements. The relative paucity of published papers interpreting and analysing its trove of data may be related to the degradation in quality of the images captured by the SODISM telescope. Some key works published on the PICARD satellite are summarized below.

Lefebvre et al. (2004) described the development of a method that uses the singular spectrum analysis (SSA) of two signals applied on MIRESOL (Mesures et Identifications des Régions actives au limbe Solaire) instruments. These have the same aperture as that in SODISM. The first signal is called the radius, and is defined as the distance between the centre and inflexion point of a limb profile, and the second signal is called intensity, which represents the intensity of signal. This method is based on the decomposition of three heliographic position angle series, where the first is called the radius, the second is called the intensity, and the third is called the errors. There are input values to the next stage of the algorithm which decomposes these three series into components for analysis. This method does not need to remove the centre-to-limb darkness to detect features (Lefebvre and Rozelot, 2004).

Meftah's 2010 study of the PICARD payload data centre and the ground instrument SODISM II advanced our knowledge of the PICARD mechanism and provided specific details about the SODISM instrument and its thermo-elastic properties. The authors also described PICARD Payload

Data Centre (PPDC). PPDC products consist of three data levels L0, L1 and L2A. Level 0 assembles the data packets incoming from the satellite, Level 1 processing sequences consist in L0 product calibration to create daily products containing the auxiliary images. L2A processing sequences consist in computing the mean solar radius from each limb image and its daily variations. Measurement accuracy and noise are also estimated.

Etcheto et al. (2011) investigated the ghosting effect in SODISM images. Advanced Systems Analysis Program (ASAP) simulation from the Breault Research Organization (BRO) to perform wave-optics calculations was used to study the effect of stray light with broad sources and high order splits. The paper discussed ghosting effects in SODISM images, and explained designs which can reduce such effects. Unfortunately, some residual ghosting could not be totally removed due to mechanical tolerances as shown in Figure 17.

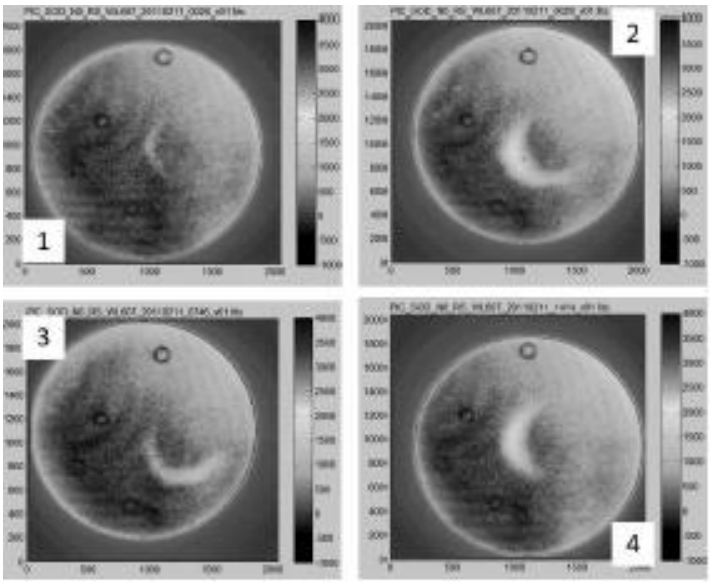


Figure 17: Images showing ghosting at wavelength 607 nm (Etcheto et al. 2011)

Qahwaji et al. (2015) designed a method for the detection of sunspots from SODISM images in the 535 nm band between 5th August 2010 and 4th January 2014, and introduced internal work[7]. The wavelet Haar filter was applied to remove noise from the images, followed by a band-pass filter to

[7] https://projects.pmodwrc.ch/solid/index.php/links/10-news-archive/31-deliverables

remove limb darkening. Gaussian smoothing was then applied to remove isolated noisy pixels, and the correlation coefficient between corresponding SOHO and SODISM images was found to be 0.98.

Meftah et al. (2016) applied the morphological top-hat operation followed by a bottom-hat operation, and a contrast enhancement technique was used before detecting sunspots using Otsu thresholding. They also focused on solar metrology and, when analysing PICARD data, could not find any direct link between solar activity and fluctuations in solar oblateness and solar radius. According to the authors, (Meftah et al., 2016), images from SODISM need to be pre-processed. Limited success in sunspot detection was achieved by applying a manual threshold to detect the sunspots. The Otsu thresholding technique was applied to SODISM images at 393 nm wavelength to detect sunspots (Kamboj and Rani, 2013) as shown in Figure 18. This is considered to be the first attempt at the segmentation of SODISM images.

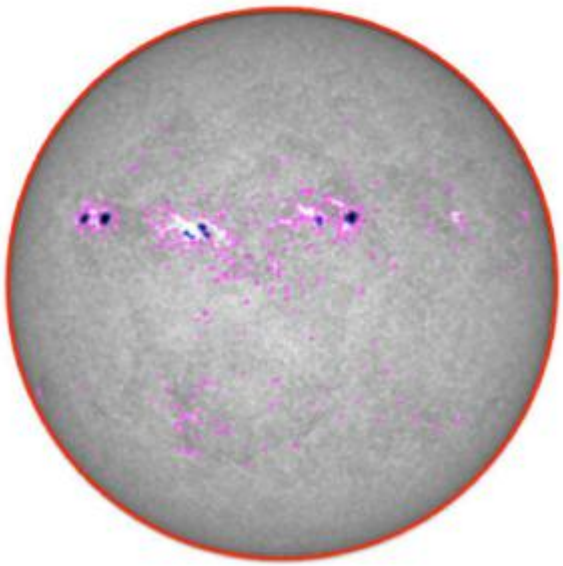


Figure 18: Disc segmentation for sunspots and faculae (Meftah et al., 2016)

Alasta et al. (2017) used an automated method to detect sunspots, which removed noise from the images by applying the Kuwahara Wavelet and á trous filters. It then applied a band-pass filter to display sunspots on a normalized background. Finally, the threshold run is obtained to mask an

image determining the sunspot location. Figure 19 shows the benchmark image and their results after applying this method.

Figure 19: A SODISM 535 image from 12:30, 1st October 2011 (left) and resulting detected sunspots (right) (Alasta et al. 2017)

2.7.2 Construction of PICARD

The PICARD satellite can be used for many purposes, such as the simultaneous measurement of spectral solar irradiance and the shape and diameter of the sun. These measurements are obtained to give a better understanding of solar activity. The PICARD payload is composed of the following instruments (Pradels, et al. 2008) :

- SOVAP (SOlar VAriability PICARD): this contains a bolometric sensor and differential radiometer to measure the total solar irradiance, which is called the solar constant.
- PREMOS (PREcision MOnitor Sensor): this includes three photometers used to measure the total solar irradiance (TSI).
- SODISM: this is an accurately pointed telescope and a CCD, which measures solar diameter and shape with the accuracy of a few milliarc-seconds. Furthermore, it observes helio-seismologic performances to analyse the solar interior. Figure 20 shows the main Stages of PICARD Project.

Figure 20: Main stages of the PICARD project[8]

Recently, a mass of solar data has been presented in visual and textual form, recorded by different observations that can be accessed online. The Virtual Solar Observatory (VSO) is a distributed data service that enables access to providers of solar data involving space and ground-based observations (Davey et al., 2004) . The public can obtain data from the following international providers: Hα Network (HANET), High Altitude Observatory (HAO), Laboratory for Space and Solar Physics (LSSP), Montana State University (MSU), Mount Wilson Solar Archive Digitization Project (MWSPADA), National Geophysical Data Center - NOAA (NGDC), National Solar Observatory (NSO), Department d'Astronomie Solaire de l'Observatoire de Paris (OBSPM), Owens Valley Radio Observatory (OVRO), Smithsonian Astrophysical Observatory (SAO), Solar Data Analysis Center - NASA/Goddard (SDAC), San Fernando Observatory (SFO), Stanford university Helioseismology Archive (SHA), and STERO Science Center – NASA/Goddard (SSC) (VSO 2005). One-minute cadence, low-resolution images (2562 pixels of about 8 × 8 arcseconds) are recorded by the SODISM telescope (Meftah and Hauchecorne, et al. 2014) on the PICARD satellite (Schmutz et al., 2009) which are available for a helio-seismology “medium-l” program in intensity (Corbard et al. 2008).

2.7.3 The SODISM interface characteristics

PICARD’s payload is managed by a global on-board electronics system known as the PGCU (PICARD gestion charge utile). The weight of SODISM is about 27.7 kg and it consumes 43.5 W of electrical power (including the PGCU). SODISM acquires a solar image every minute, and processes it on-

[8] https://microcarb.cnes.fr/fr/MICROCARB/Fr/GP_mission.htm

board. It is then transmitted as compressed data, since the global rate is about 2.2 GB per day (Meftah et al., 2010).

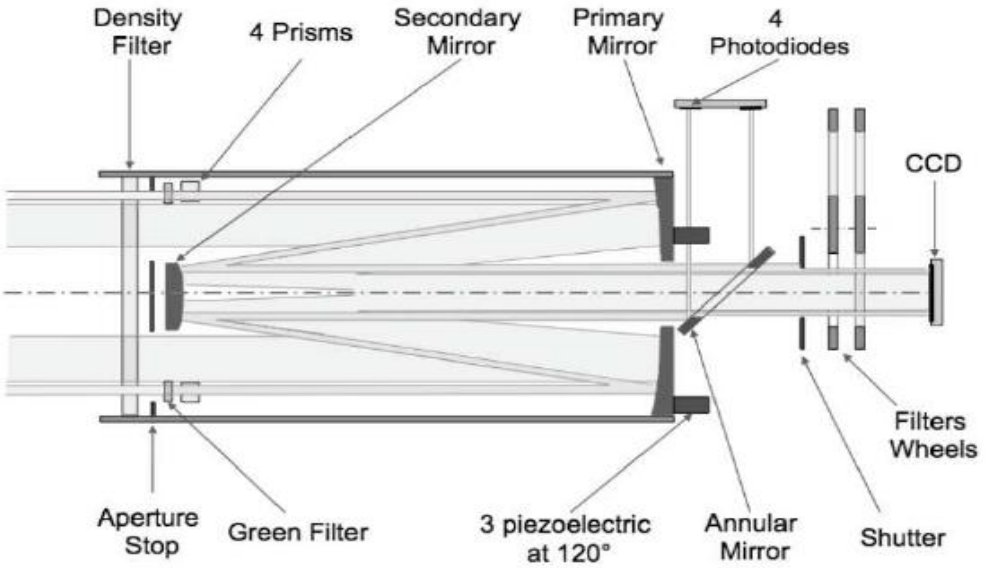


Figure 21: SODISM instrument components (Meftah et al., 2011)

2.7.4 Scientific domains addressed by SODISM

The mission of the SODISM telescope includes investigating many scientific questions about solar physics and Sun-Earth relations (Meftah et al., 2018), including:

- Measurement of the solar diameter and limb shape in the continuum wavelength, referring to the heliographic latitude.
- Detecting the active areas, such as faculae and sunspots, that may damage the diameter measurements.
- Studying, in collaboration with PREMOS, the effects of solar activity, solar radiance images and solar spectral irradiance.
- Providing information and data on the Earth's atmospheric properties, including refraction and absorption by atmospheric constituents, via the period of solar occultation.

2.7.5 SODISM solar data

In this research, large data sets were downloaded from the PICARD website from August 2010 to January 2014.

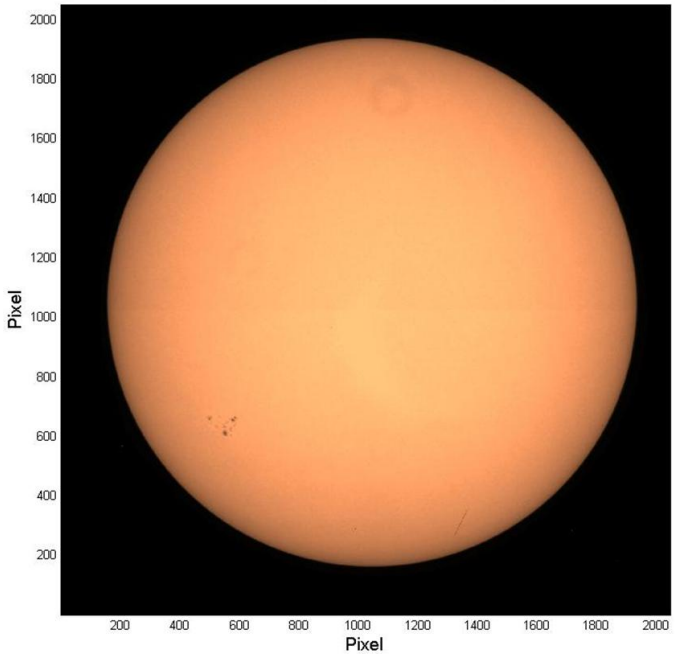


Figure 22: First image of the Sun captured by the SODISM instrument on 22nd July 2010 at wavelength 607 nm (Meftah et al., 2016)

SODISM uses four spectral bands to measure solar irradiation, and for calibration there are two bands (flat field and star field). To avoid the presence of Fraunhofer lines, the application of a narrow spectral domain was used in the solar continuum to measure the diameter through three filters, centred at 535, 607 and 782 nm (Rouzé et al., 2014).

Table 1 illustrates SODISM instrument detection characteristics where the first column shows the wavelengths of SODISM, the second shows Δλ of each wavelength and the third column shows the function performed by each band.

Table 1: SODISM instrument detection characteristics (Rouzé et al., 2014)

λ (nm)	Δλ (nm)	Function
215	7	Solar activity, O3 measurement, diameter
393.37	0.7	Observation of active regions
535.7	0.5	Diameter, (helio-seismological oscillations)
535.7	0.5	Diameter
607	0.8	Diameter
782.2	1.6	Diameter
FOV (Field of View)		35 arcmin
Angular resolution		1.06 arcsec
Measurement precision		< 3 marcsec (mas) per image
Flux resolution		< 10-5/pixel

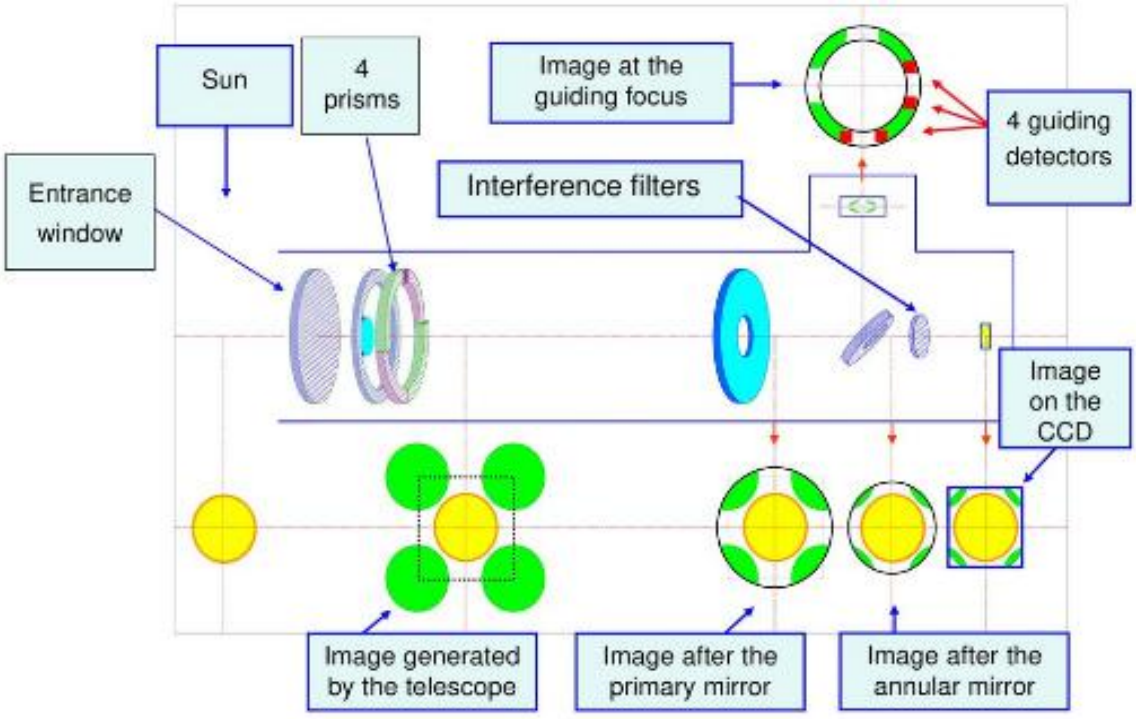


Figure 23: Block diagram of the SODISM optics scheme (source: CNRS/SA)[9]

[9] https://directory.eoportal.org/web/eoportal/satellite-missions/p/PICARD

The operator of the PICARD payload data centre takes charge twice a week for nominal programming; this post also involves the mission programming responsibility on the PICARD payload data centre to determine measurement dates (Meftah et al., 2016).

2.7.5.1 SODISM programming

SODISM is a telescope using 2k×2k pixels charge couple devices (CCD) (Pradels et al., 2008.), and a 10^{-3} arc-second relative accuracy of measurement is used. There are various different types of measurement, each defined by a word of 2 bytes, that indicates the wavelength (dark current 215 nm 393 nm 535 nm 607 nm and 782 nm), on-board processing (software compression), and image characteristics, such as size (narrow/wide limb, full/window images), creating a table of around 64 types. However, only a small part was used through the nominal mission. There is a 22 pixel-wide strip of solar limb recorded at 535 nm every 120s (helio-seismology); meanwhile a 40 pixel-wide limb strip is recorded at 215 nm, 393 nm, 535 nm, 607 nm and 782 nm, with a cadence of two of them per orbit, while the others are recorded each minute at 535.7 nm. Figure 24 illustrates the image types from the SODISM telescope (Meftah et al., 2010).

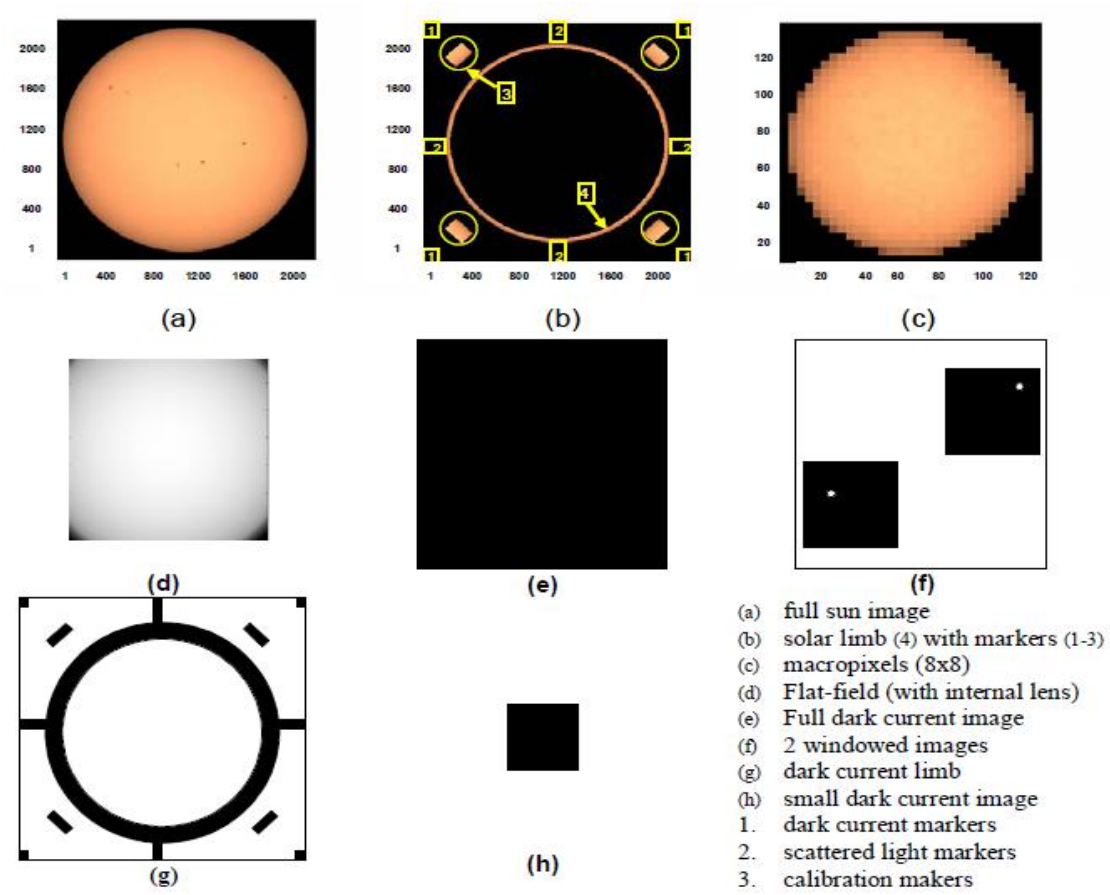


Figure 24: Image types from the SODISM telescope (Meftaha, et.al. 2010)

The dark signal limb images are 40 pixels wide, and not specific for any wavelength. Figure 25 shows some markers which are pieces of images extracted from the SODISM CCD image and 4 auxiliary ones, which are present with the limbs, these limb images registered at the wavelength at 535.7 nm.

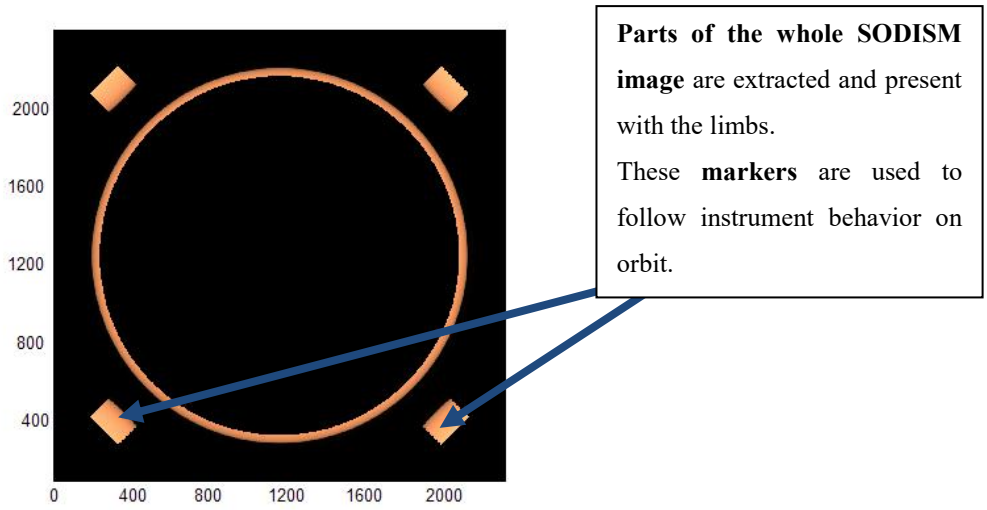


Figure 25: Extract of SODISM image with limb

Two types of compression are applied on the PICARD board, given the limited volume of telemetry per day:

- Centre National d'Etudes Spatiales, Edouard (CNES) provides loss-less compression called the PICARD loss-less cruncher (PLLC).

- Y. Langevin designed a compression with a loss rate of 16 (Assus et al., 2008). This rate is applied to prevent artefact effects in sunspot locations used to retrieve the Sun's different rotations.

The other types of images used for nominal configuration are listed as follows (Pradels, et al., 2008):

[i] Limbs: 22 or 40 pixel width compressed using PLLC algorithm.

[ii] Full images: 2k×2k pixels, compressed using the Langevin algorithm.

[iii] Macro-panelised images: averaging over 8×8 pixels compressed using PLLC.

[iv] Windowed images: 768×768 pixels compressed using the PLLC algorithm.

[v] Small images: 256×256 pixels with no compression.

All image types can be recorded whatever their wavelength filter. A dark image will be recorded once the shutter of an instrument remained closed.

Two images will be obtained from one measurement on board: the full solar image and the corresponding solar limb. There are two modes in which the SODISM instrument can operate:

1. Measurement mode
2. Investigation mode.

The measurement mode is used for normal scientific operations, while investigation mode is used for manually operated checks (Pradels et al., 2008). Table 2 illustrates the number of SODISM images per scenario/day.

Table 2: The numbers of images foreseen in the SODISM mission scenario (Meftah et al., 2016)

Image Type	Mission Mode	IMG/DAY
Macropixel 8x8 pixels, 535 nm	Nominal	1235
	Night	294
Limb, 20 pixels, 535 nm	Nominal	720
	Night	294
Limb, 40 pixels, dark current	Nominal	58
	Optical distortion	60
	Night	210
Limb 40 pixels, 215, 393, 535, 607, 782 nm	Nominal	145
	Optical distortion	600
	Night	135
Full images 2048x2048 pixels, 393 nm	Nominal	14
Full images 2048x2048 pixels, 215, 393, 535, 607, 782 nm	Nominal	14
Full images 2048x2048 pixels, dark current	Nominal	1
	Dark current 1	100
Full flat-field images 2048x2048 pixels, 215, 393, 535, 607, 782 nm	Nominal	1
Windowed images 786x786 pixels	Stellar	20
Small images 256x256 pixels	Dark current 2	1440
Total by scenario/day		
Nominal		2188
Optical distortion		660
Night		798
Absorption		135
Dark current 1		100
Dark Current 2		1440

Table 3 shows the data daily rates of images captured by the SODISM and varying in wavelength.

Table 3: Daily data rate for the SODISM

Wavelength λ in nm	Daily data rate (nominal mode at beginning of the mission)
215	3-4
393.37	11
535.7	3-4
535.7	3-4
607.1	3-4
782.2	3-4

2.7.5.2 Solar activity catalogues

Several satellite and ground-based stations study the Sun and provide their observations as visual images or in textual form. The data for this research was obtained from the PICARD website[10] which provides SODISM images. Full images at 2048×2048 pixels were acquired for the following wavelengths: 215 nm, 393 nm, 535 nm, 607 nm and 782 nm

2.7.5.3 Characteristics of SODISM data

There are three main characteristics (Corbard, T., et al. 2013) of the data;

- Regular passages through the South Atlantic Anomaly (SAA) due to the low PICARD orbit (Irbah et al., 2012). In a 24-hour period, the SAA corresponds to about 7% of the measurements.
- An orbital period around the Earth of about 100 minutes produces aliased peaks in the power spectra of the low-degree p-mode signals.
- The presence of CCD affects about 10% of the images and the SODISM intensity signal has a lower SNR compared to SDO HMI.

[10] http://PICARD.busoc.be/sitools/invoquerSva.do?sva=PICARD&svaAction=searchFiles,

Figure 26 illustrates an additive signal in a SODISM pixel; the biggest challenge is to discover the features that can be detected in each band.

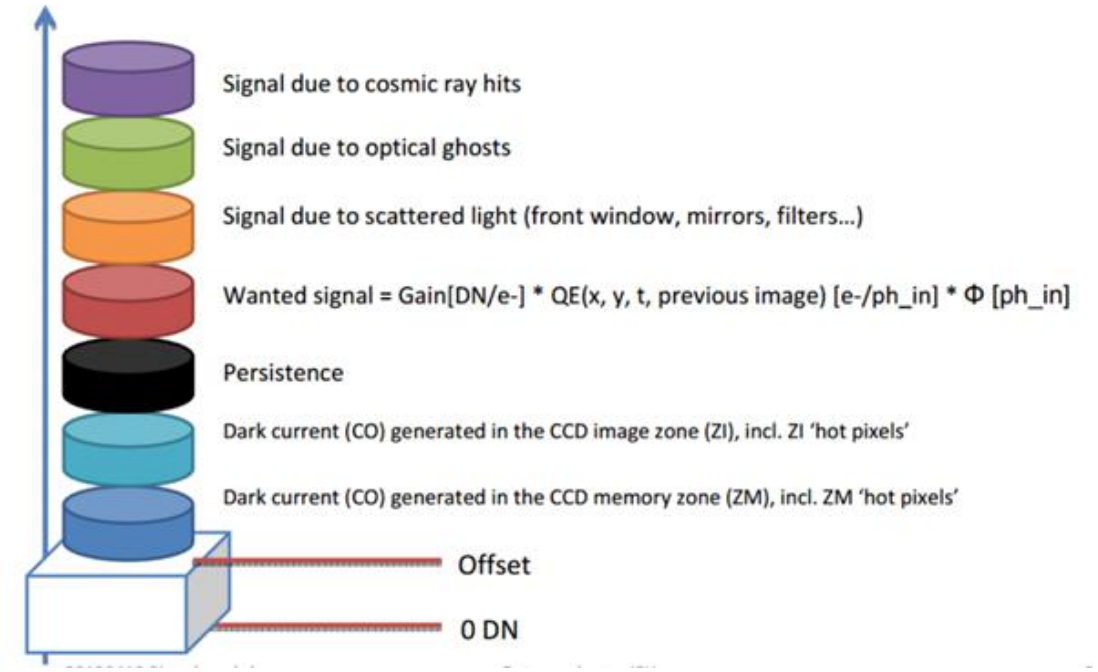


Figure 26: Additive signals contributing to a SODISM pixel (Hochedez, 2012)

2.8 Data challenges and limitations of SODISM

The greatest challenge is producing an efficient methodology to remove noise from the images. Noise removal is a major concern in computer vision and image processing, because noise can cause image blurring and degradation (Kaur and Shukla, 2014). Other challenges tackled in this thesis relate to image enhancement and the accurate detection of solar features. In addition, solar features can be detected at different wavelengths, and there is a need to associate segmentation results from different wavelengths to establish an indication of accuracy. Moreover, access to SODISM data became a significant concern, and subsequently a problem, in the course of this research, as public access to the data was restricted halfway through this study. The final challenge encountered while undertaking this research related to the lack of published research on SODISM image analysis.

2.9 Solar spectral irradiance (SSI)

Some of the images from SODISM suffer from poor quality. The effects on the SODISM space-based telescope are detailed in the following points (Meftah et al. 2016):

i. Additive (such as dark signals, offset, cosmic ray hits, ghost images).
ii. Multiplicative (such as nonlinearity of detection, optical flat field, charge-coupled device (CCD) flat field).
iii. Convolutive (such as PSF scattered light, kinematic blur, optical aberrations, persistence/hysteresis, CCD charge).
iv. Other effects, such as distortion (anamorphosis).

2.10 Conclusions

Publications discussing the enhancement and segmentation of SODISM images are few in number. SODISM images need further investigation to achieve a better understanding of how to design more reliable automated sunspot detection systems, and to overcome the degradation in their quality. Previous studies illustrate many challenges, summarized as follows:

- There is a lack of research into SODISM enhancements and segmented images, which makes evaluation difficult but not impossible. Any new system of SODISM enhancement or segmentation could be compared with other satellite image catalogues, such as SOHO, to check the accuracy and type of assessment that can be used to compare manually and automatically detected sunspots. This can also be checked against the National Oceanic and Atmospheric Administration's (NOAA) catalogues (Qahwaji and Colak, 2005a).

- The automated system must ensure that the collection of data and SODISM image wavelengths are up to date, or risk the segmentation of sunspots. The 2016 paper of Meftah et al. applied segments on 393 nm but did not do so to all the images. Thus, their work is incomplete and is still in progress (Meftah et al., 2016).

- The reference images are not available for SODISM, so to achieve the goals, quantitative assessment approaches using mean square errors (MSE) have been applied, moreover, Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM) are calculated to evaluate image enhancement quality. Furthermore, Normalized Absolute Error (NAE), Structural Content (SC) and Normalized Cross-Correlation (NCC) have been calculated to assessment enhancement method.

- The life period of PICARD where full data can be obtained (August 2010 - Jan 2014) is limited; furthermore, this period is missing a lot of data. For example, whilst there is data for W.L 215 nm in 2012, 2013 and 2014, there is a lack of quality in the images, which suffer from noise and unwanted features that need pre-processing before segmentation.

- Whilst the SODISM telescope provides important data, these still need to be analysed to produce information that can be used for different systems in the future, such as automated sunspot detection or to develop and assess flare prediction models.

CHAPTER THREE

3.0 Practical implementation and an evaluation.

3.1 Introduction

This chapter presents two major procedures, namely a technique for image quality enhancement, and an algorithm for segmenting SODISM images. To design and evaluate the algorithms, SODISM images collected between August 2010 and January 2014 are utilised to achieve quality enhancement, and the DWT is used to preserve the high-frequency components of the image features. More specifically, the corresponding high-frequency sub-bands of the IDWT are emphasized so that edge features are enhanced in the images, and this process is summarised in the algorithm shown in Figure 27 (Alptekin Temizel, 2007).

This chapter is organized as follows. Section 3.2 discusses the analysis of data, including the pre-processing of the data, and Section 3.3 proposes the quality enhancement method. Following on, the assessment and evaluation procedure is outlined in Section 3.4, and Section 3.5 describes the feature extraction techniques. The results of the image enhancement technique evaluation are presented in Section 3.6, and in Section 3.7, the image segmentation is presented in detail. Finally, the conclusions and a summary of the findings are presented in Section 3.8.

3.2 Data analysis

Several research areas are dedicated to clarifying the analysis of solar image segmentation and feature detection, in particular solar irradiance reconstruction and solar activity prediction. On a practical basis, image segmentation is often associated with the uncertainties which can result from noise distortion data, feature definition, and the identification process. Therefore, an initial vital step to be considered before segmentation is data pre-processing.

The data utilised in this work involve two types of images obtained from the PICARD satellite. The first image type is provided in the FITS format at a wavelength of 535 nm, whilst the second type is in a JPG format. Due to the

inherent data distortion issues associated with the JPGs format, this type needs an enhancement procedure. Choosing a suitable approach for isolating different objects from each other, or from their background, is considered the most common challenge. Several methods of isolating the different features are available, raising the question as to which is the most suitable for these images. Various types of images have a different characteristic, which means that every set of data needs a distinct approach for its analysis and segmentation. The primary step in object detection, recognition and identification in image analysis is segmentation.

SODISM data available online was downloaded[11]. The date chosen was 25 December 2010 for five different wavelengths of 215 nm, 393 nm, 535 nm, 607 nm and 782 nm, respectively. Practically, the data represents solar images at different times. However, approximately the same time was chosen for the different wavelengths in analysing the images using the same technique. Moreover, three different times for each wavelength were chosen.

3.2.1 Data pre-processing

The usual way to accurately detect whole sunspots is by using an iterative detection method that involves the periodic increase of the scope, scale and counting the detected pixels (Curto et al. 2008). The first kind of data used in this thesis were image files for five wavelengths (215 nm, 393 nm, 535 nm, 607 nm, and 782 nm) in JPG format, where each image has a size of 2048 × 2048 pixels. As previously stated, the data were collected over a period of four years (from 5th August 2010 to 4th January 2014). It is worth mentioning that these data are no longer available on the PICARD satellite website. The poor quality of these images necessitates image quality enhancement, leading to the development of a quality enhancement algorithm. This enhancement pipeline selects an image for quality enhancement and then feeds it into the segmentation algorithm.

[11] http://PICARD.busoc.be/sitools/invoquerSva.do?sva=PICARD&svaAction=searchFiles

3.3 Proposed image enhancement method

The improvements in the quality of SODISM images are undertaken in the wavelet domain. Initially, wavelet coefficients are estimated, which provides an opportunity to improve the sharpness and edge features of the resulting inverse transformed image. Since the DWT shatters the data into more components (time and frequency), it becomes easier to filter in or out a given non-stationary waveform. Moreover, the discrete wavelet transform method has been found, both visually and by using quantitative measures, to give better results than interpolation techniques. This includes FFs that are calculated for solar features (Ashamari, Qahwaji, Ipson, Schöll, et al., 2015) (Piao et al., 2007). Many conventional methods for image resolution enhancement, such as bicubic and bilinear interpolation approaches, can generate blurred images and provide false information, and this is due to the use of inaccurate information for the analysis that is relevant to the edges in the original image (Fowler, 2005).

This work explores the application of size increasing techniques to SODISM in order to improve the visibility and measurements of solar features, such as sunspots. The main deficiency after the application of interpolation techniques is the smoothing caused by interpolation (Kumar et al., 2013b). Wavelet-based methods can provide a way of smoothing, by enhancing image resolution through estimating the high-frequency information from the given image. Such methods are based on the idea that the image to be enhanced is a low-frequency sub-band among wavelet-transformed sub-bands, and the aim is to estimate the corresponding high-frequency sub-bands, so that an inverse wavelet transform can then be performed to obtain an image with enhanced resolution. The input image to be enhanced is regarded as a low-frequency sub-band in the context of an IDWT.

The down-sampling of the DWT sub-bands causes a loss of information in the respective sub-bands, and the SWT is used to reduce this loss. Both the SWT high-frequency and interpolated high-frequency sub-bands are the same size, which means they can be added to each other. The new corrected high frequency sub-bands can be interpolated further for greater

enlargement, and the low resolution image is obtained by the low-pass filtering of the high resolution image (Atkins et al., 2001). Actually, a low-frequency sub-band is the low resolution of the benchmarked image. A DWT and SWT are used to decompose the input image into four sub-bands. The LL sub-bands are rejected and the pairs of SWT and DWT low-high (LH), high-low (HL) and high-high (HH) sub-bands are combined after the latter have been increased in size to match the former via bicubic interpolation. Assuming that the output image is to be interpolated by a factor α, (α =2) over the input image, the high-frequency sub-band images are interpolated by a factor α and the input low resolution image is interpolated by a factor α before they are input to the IDWT to produce the enhanced image.

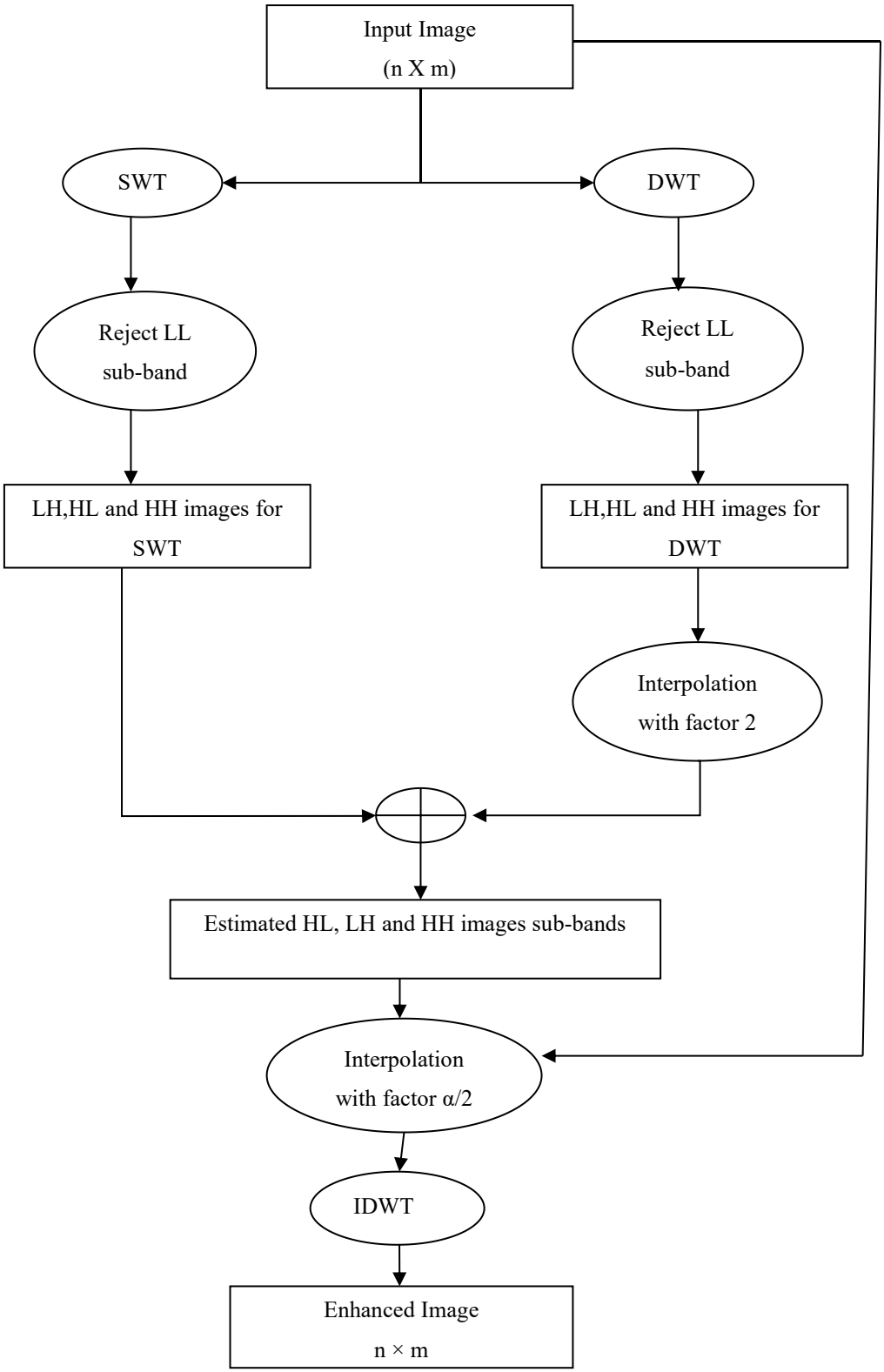


Figure 27: Algorithm to obtain an image enhanced by the proposed technique

3.4 Assessment procedure

The proposed technique has been tested on many SODISM satellite benchmark images taken on the same day at different times and at different wavelengths. The test image subjects are obtained by reducing the resolution of the original benchmark SODISM images by applying a DWT, and preserving the LL sub-band, as illustrated by Demirel and Anbarjafari's study (2011). Two approaches are used to examine the effectiveness of the enhancement methods used. The first is the visual comparison of the benchmark (original) and enhanced images. The second, for a quantitative comparison, is a calculation of the peak signal-to-noise ratio (PSNR), using the following formula:

$$\mathrm{PSNR} = 10\log\left(\frac{\mathrm{R}^2}{\mathrm{MSE}}\right) \quad (9)$$

Here R is the maximum fluctuation in the input image (255 in this study), MSE represents the Mean Squared Error between the benchmark image, I_{in}, compared with the enhanced image, I_{out} is defined as follows,

$$\mathrm{MSE} = \frac{\sum_{\mathrm{i,j}} (\mathrm{I_{in}(i,j)} - \mathrm{I_{out}(i,j)})^2}{\mathrm{M}\times\mathrm{N}} \quad (10)$$

where *M* and *N* are the dimensions of both the benchmark and enhanced images (2048 × 2048 pixels). Both the MSE and PSNR measure the similarity of the two images based on pixel-by-pixel matching. The MSE value decreases to zero while the PSNR increases without limit. Although the MSE and PSNR provide quantitative measures of similarity between pairs of images, they do not correlate highly with human visual perception. Therefore, more quantitative comparison and visual quality assessment measures are also needed to monitor the quality of SODISM images. The proposed method has been applied for seven days per month in the year for five wavelengths, except for the 215 nm wavelengths where data are not available for 2012 and 2013. Since the image features detected at various wavelengths differ, this method of computing peak signal-to-noise ratios is the most appropriate method of evaluation. Afterwards, three different techniques have been compared: bicubic interpolation, wavelet zero padding (WZP) and the proposed discrete wavelet transform (DWT). Moreover, some

of the results of the image enhancement procedure are shown in Figures 28, and 29.

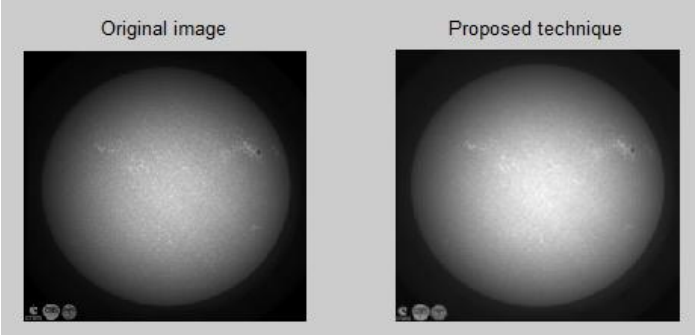


Figure 28: Original Image and image after processing at wavelength 215 nm at 23:15. PSNR for proposed technique = 39.011

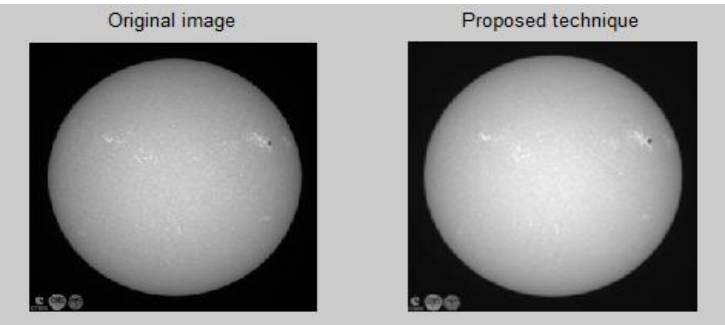


Figure 29: Original image and Image after processing wavelength 393 nm and at time 23:17. PSNR for proposed technique = 3643×10^{-2}

Table 4: Different wavelength images results for different PSNR

Wavelength	Time	Peak Signal to Noise Ratio			%
		Bicubic	WZP	DWT	DWT:BIC
215 nm	04:05	3452×10^{-2}	4041×10^{-2}	4058×10^{-2}	14.91%
	09:45	3786×10^{-2}	4391×10^{-2}	4390×10^{-2}	13.76%
	10:53	3764×10^{-2}	4350×10^{-2}	4369×10^{-2}	13.83%
393 nm	04:07	2967×10^{-2}	3559×10^{-2}	3560×10^{-2}	16.64%
	16:21	3083×10^{-2}	3673×10^{-2}	3675×10^{-2}	16.11%
	09:47	2967×10^{-2}	3558×10^{-2}	3561×10^{-2}	16.64%
535 nm	04:43	3015×10^{-2}	3600×10^{-2}	3611×10^{-2}	16.47%
	12:13	2998×10^{-2}	3590×10^{-2}	3595×10^{-2}	16.55%
607 nm	07:15	3056×10^{-2}	3646×10^{-2}	3649×10^{-2}	16.23%
	14:01	2996×10^{-2}	3588×10^{-2}	3590×10^{-2}	16.51%
	21:09	2871×10^{-2}	3460×10^{-2}	3466×10^{-2}	17.15%
782 nm	02:53	2859×10^{-2}	3449×10^{-2}	3452×10^{-2}	17.17%
	08:09	2875×10^{-2}	3464×10^{-2}	3467×10^{-2}	17.09%
	16:13	2838×10^{-2}	3428×10^{-2}	3432×10^{-2}	17.29%

The results for the first attempt are summarized in Table 4 for three different times at the five wavelengths of 215, 393, 535, 607 and 782 nm, which appear in the column to the left. The figures obtained for the PNSR using the three methods are shown in the remaining columns. The initial results show greater improvements in the definition and clarity of images through using the DWT method for all SODISM image wavelengths compared with other methods. Table 4 shows improvement in all wavelengths and the average improvement is approximately 16%. Further discussions of the results are presented in the rest of this chapter.

3.6 Experimental results

As previously mentioned, SODISM images for the five wavelengths 215, 393, 535, 607 and 782 nm were downloaded[12] for the first week of each year from 2010 to 2014 (when available) giving approximately 900 images in JPG format. The results are shown in Figure 99 and in Tables 5 and 6. Figures 30, 32, 34, 36 and 38 show the results for 1st October 2011, near the beginning of the satellite's operation. The downloaded images are used as the benchmark images and, using a value of **2** for the factor '*α*', test images were constructed from them, as indicated in Section 3.3. The DWT used throughout this work was the Daubechies 9/7 wavelet transform, which was employed to decompose the input images into different sub-bands (Meftah et al., 2015).

Although the SODISM images used were taken at different times, they represent the closest times available for the different wavelengths by which they could be processed using the same technique. Thus, in Figures 30 to 38, the features visible in the SODISM images vary from one wavelength to another, and from Figure 40 onwards, the images vary with time. These variations are due to out-of-focus optics, motion blur, and the deteriorating conditions of the instrument.

[12] http://PICARD.busoc.be/sitools/invoquerSva.do?sva=PICARD&svaAction=searchFiles

3.6.1 Experimental set-up

All experiments were conducted using a PC with a Core i5 2.5 GHz CPU processor and 6 GB of RAM, running Windows 7 Home Premium edition and MATLAB_R2012b software.

3.6.2. Experimental details

For evaluation, MSE values are used to compare the benchmark images with the enhanced versions of the test images shown in Figure 102 and Table 5. Meanwhile, the PSNR values are shown in Table 6. As can be seen in Tables 5 and 6, the proposed technique illustrates a nearly fourfold improvement in the MSE values, and an improvement of nearly 6 dB in the PSNR values in comparison to the bicubic interpolation for all five wavelengths. Visual comparisons of the test images with their corresponding enhanced images are shown in Figures 30 to 39. These all show the excellence of the proposed image quality enhancement technique and its superiority over bicubic interpolation. A further advantage is that the bicubic interpolation processing took 40 seconds to complete, while the proposed technique took only 10`. Furthermore, magnified local features were used to visually compare the SODISM images before and after the enhancement, as shown in Figures 31, 33, 35, 37 and 39. The improved visibility of sunspot regions in the enhanced images compared with the corresponding test images can be clearly observed.

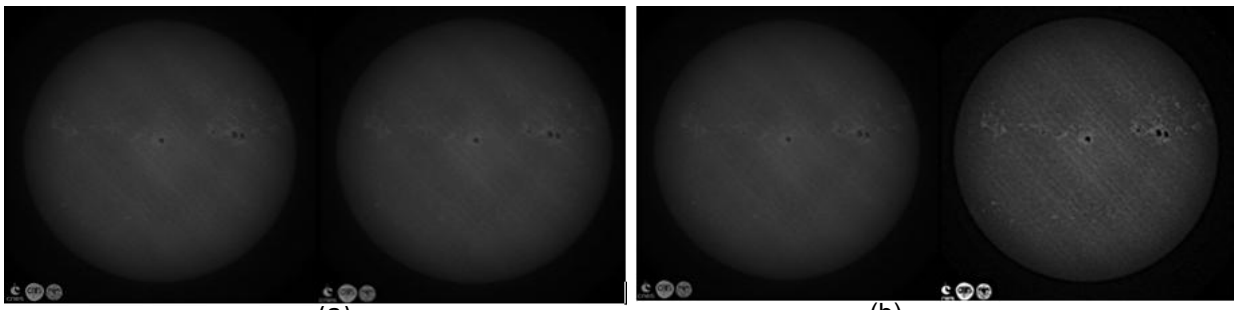
Table 5: Comparison of MSE values obtained by bicubic interpolation and the proposed technique for different wavelengths

Wavelength and Time	MSE of Bicubic interpolation	MSE of Proposed technique
215 nm at 02:55	1.9×10^{-5}	5×10^{-3}
393 nm at 02:47	97×10^{-3}	24×10^{-3}
535 nm at 12:30	192×10^{-3}	49×10^{-3}
607 nm at 01:01	353×10^{-3}	89×10^{-3}
782 nm at 02:07	309×10^{-3}	79×10^{-3}

Table 6 shows that a comparison between bicubic technique and proposal technique values of PSNR in all SODIUM wavelengths is improved by approximately 12%.

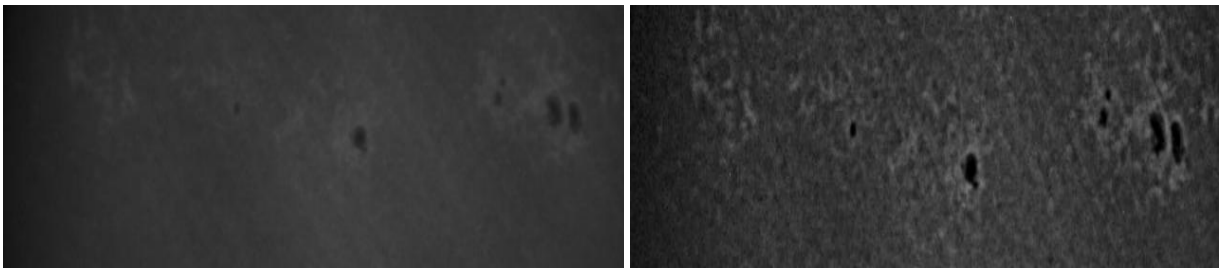
Table 6: Comparison of PSNR values obtained by bicubic interpolation and the proposed technique for different wavelengths

Wavelength and Time	Peak Signal-To-Noise Ratio (PSNR)		
	Bicubic interpolation	WZP	Proposed technique
215 nm at 02:55	4115×10^{-2}	4703×10^{-2}	4710×10^{-2}
393 nm at 02:47	3421×10^{-2}	4021×10^{-2}	4025×10^{-2}
535 nm at 12:30	3124×10^{-2}	3522×10^{-2}	3717×10^{-2}
607 nm at 01:01	2859×10^{-2}	3452×10^{-2}	3454×10^{-2}
782 nm at 02:07	2916×10^{-2}	3504×10^{-2}	3509×10^{-2}



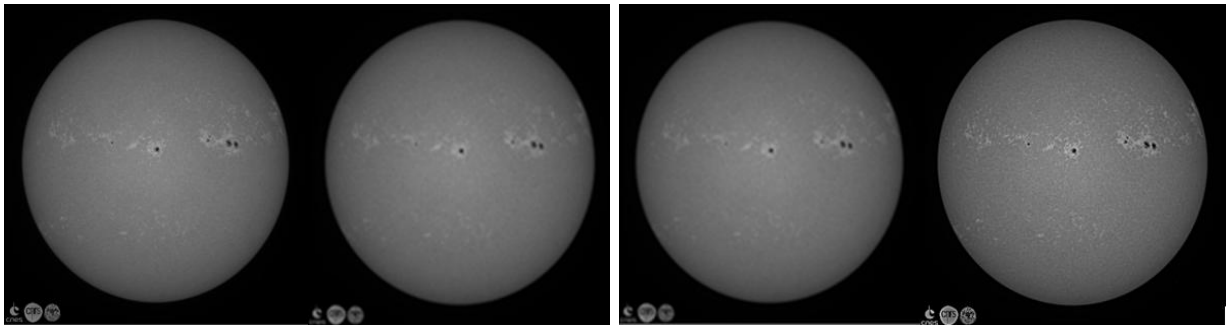
(a) (b)

Figure 30: (a) Test image on the left and image enhanced by bicubic interpolation on the right: (b) Reference image on left and proposed technique on right for wavelength 215 nm at 02:55



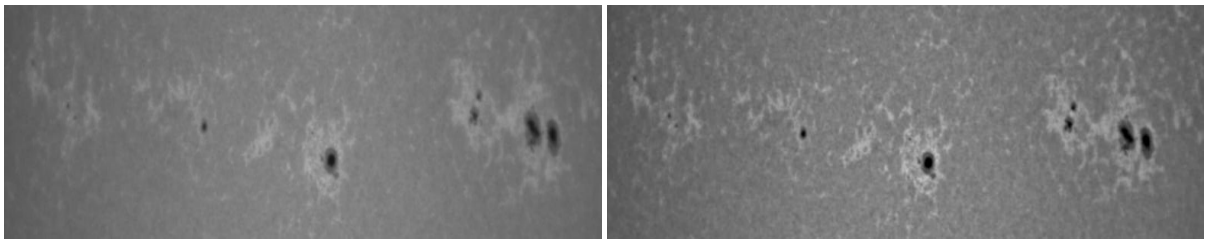
(a) (b)

Figure 31: (a) Magnified sunspot region for test image of figure 30 (a) and (b) After enhancement using the proposed technique.



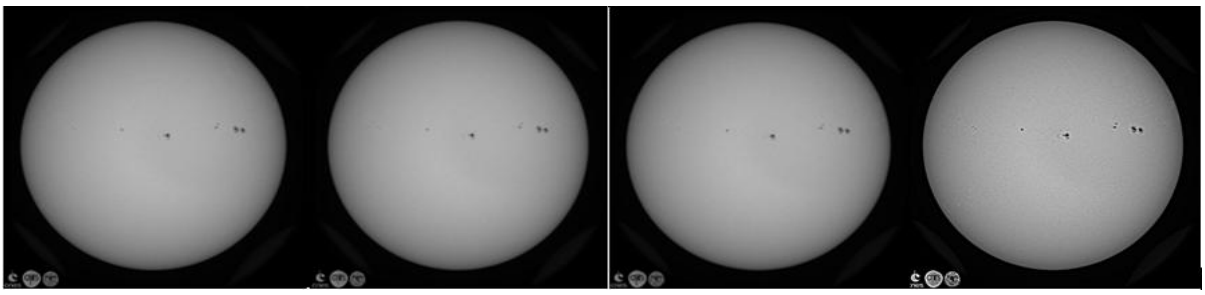
(a) (b)

Figure 32: Figure 32: (a) Test image on the left and image enhanced by bicubic interpolation on the right. (b) Reference image on left and after use of proposed technique on right for wavelength of 393 nm at 02:47



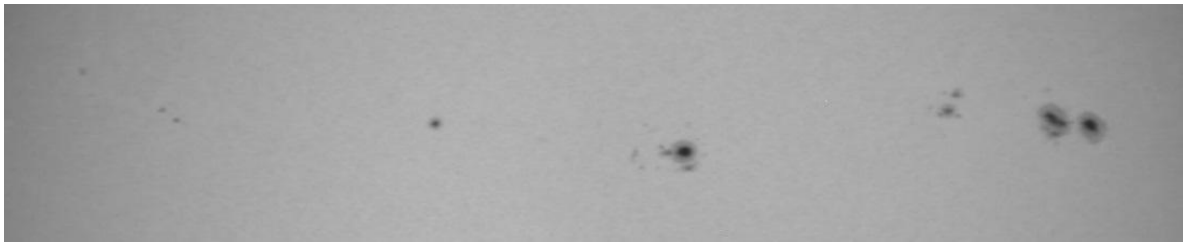
(a) (b)

Figure 33: (a) Magnified sunspot region for test image of figure 32(a) and (b) After enhancement using the proposed technique

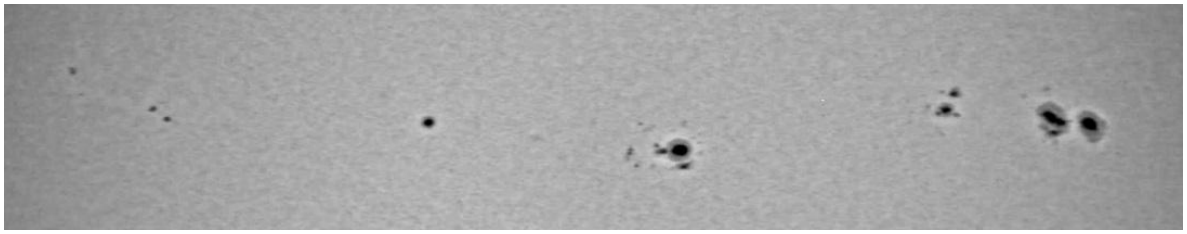


(a) (b)

Figure 34: Figure 34: (a) Test image on left and image enhanced by bicubic interpolation on the right, (b) Reference image on the left and after using the proposed technique on the right for wavelength of 535 nm at 12:47

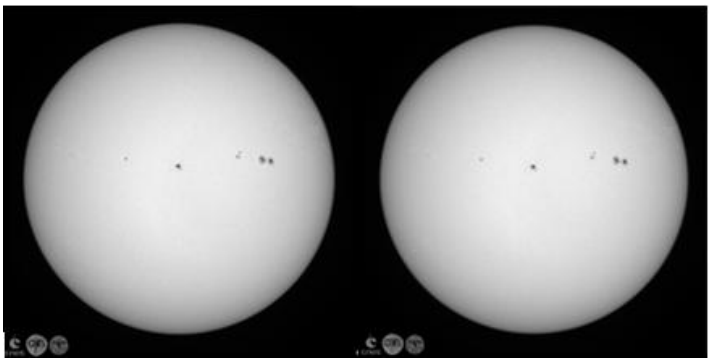


(a)

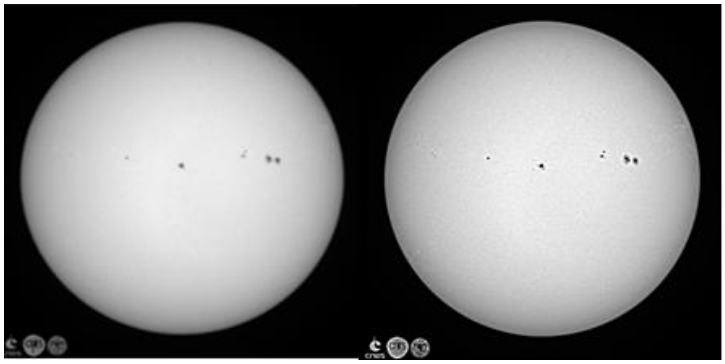


(b)

Figure 35: (a) Magnified sunspot region in test image of figure 34(a) and (b) After enhancement using the proposed technique

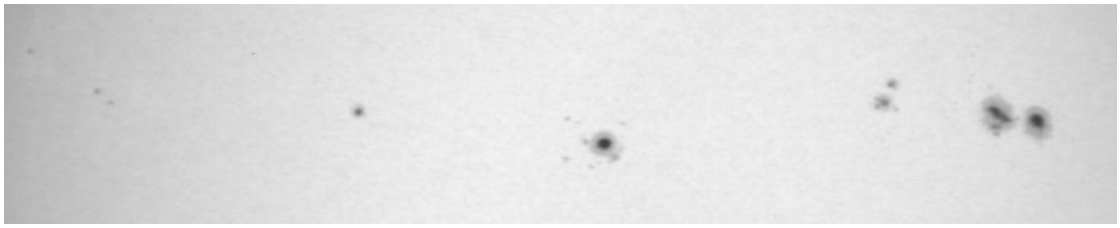


(a)

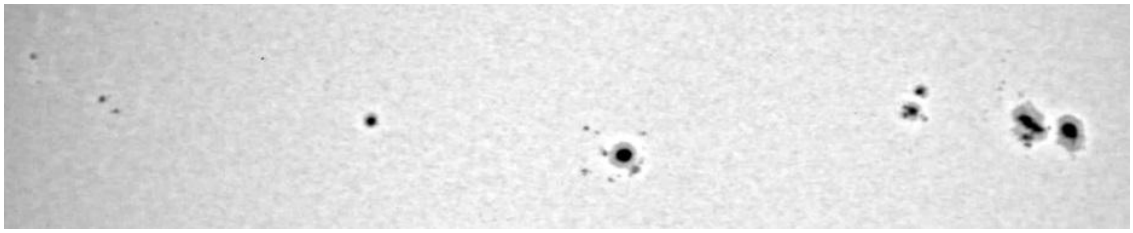


(b)

Figure 36: (a) Test image on the left and image enhanced by bicubic interpolation on the right, (b) Reference image on the left and after using the proposed technique on the right for wavelength of 607 nm at 01:01

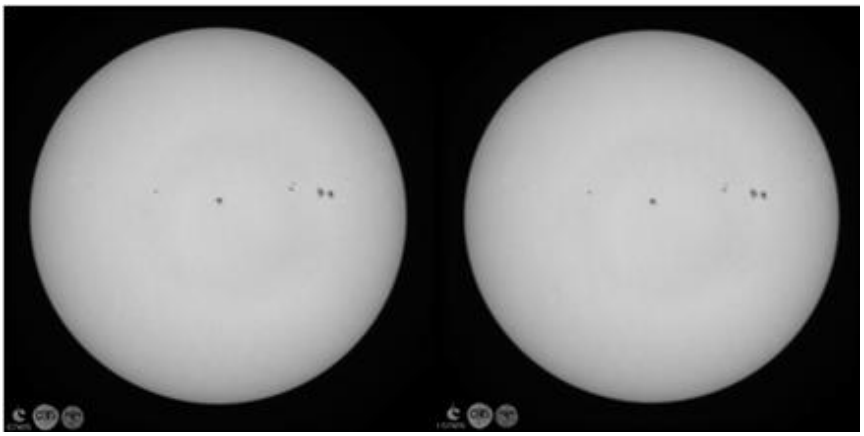


(a)

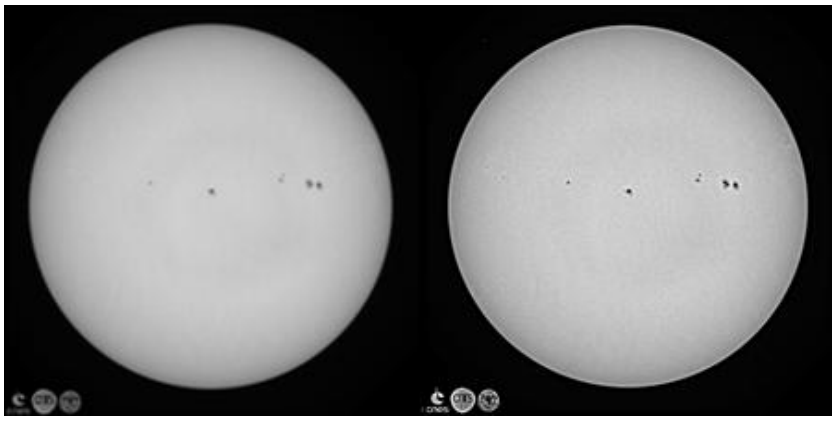


(b)

Figure 37: (a) Magnified sunspot region in test image of figure 36(a) and (b) After enhancement using the proposed technique

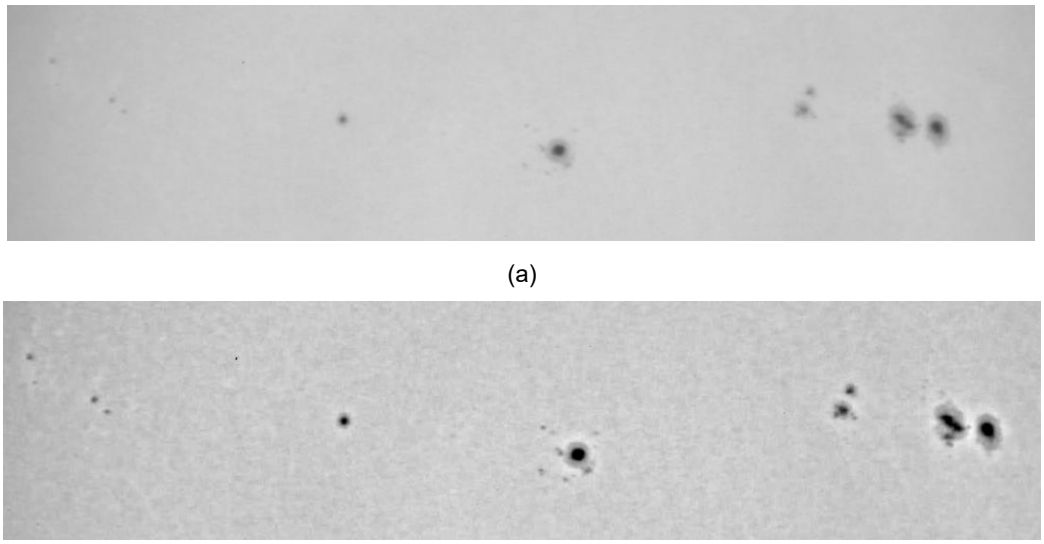


(a)



(b)

Figure 38: (a) Test image on the left and image enhanced by bicubic interpolation on the right, (b) Reference image on the left after use of proposed technique on the right for wavelength of 782 nm at 02:07



(a)

(b)

Figure 39: (a) Magnified sunspot region in test image of figure 38(a) and (b) After enhancement using the proposed technique

3.7 SODISM image segmentation

Two segmentation experiments were conducted; the first was the segmentation of the JPG images, while the second was for the FITS images. Finally, the results were compared with those from the SOHO satellite collected for the same time period. Moreover, the findings including numbers, location, and size of the sunspots were also compared with the NOAA catalogue on the same date and approximately the same time.

3.7.1 Segmentation of JPG images

These data were downloaded from the PICARD website[13]. The central wavelengths of the five spectral bands were: 215.0, 393.37, 535.7,607.1, and 782.2 nm, usually abbreviated as `215', `393', `535', `607',and `782' (Meftah et al., 2014). Images for five different wavelengths were downloaded (215 nm, 393nm, 535 nm, 607 nm, and 782 nm) in a JPG format, and each image is sized at 2048 × 2048 pixels. The images obtained are from 5th August 2010 to 4th January 2014. In total, 900 images were collected, and the DWT image enhancement procedure was first used to improve their quality.

[13] http://PICARD.busoc.be/sitools/index.jsp

Following this step, the segmentation procedures were applied, as detailed in the following sections.

3.7.1.1 Wavelength 215 nm

The quality of the images obtained at the wavelength 215 nm was the worst, as they were darker due to the degradation of intensity during the PICARD mission. The quality of the data progressively deteriorated, particularly at 215 nm, as shown in Figure 40. After the launch of PICARD, the SODISM telescope started recording images continuously. Level 0 (L0) data were generated every day by the PICARD Payload Data Centre, and registered at BUSOC in Brussels (Meftah et al., 2013). Figure 40 shows a sample of level 1 (L1) images from August 2010 and February 2013. The 2013 data image exhibits a flatter centre-to-limb variation (CLV) for wavelengths 215 and 393 nm, caused by a loss of sensitivity in the ultraviolet (UV) range. The original 215 nm image has a prominent band oriented over the solar disc, and the spot artefacts are visible at the wavelength of 393 nm. Applying a Canny filter with Gaussian smoothing for a standard deviation of 2.14 pixels, and upper and lower thresholds to the original image of 5 and 0 respectively, results in an edge map in which the limb can be highlighted. Edges are distributed over the whole image and the image quality is too poor for processing to fit an ellipse to this data.

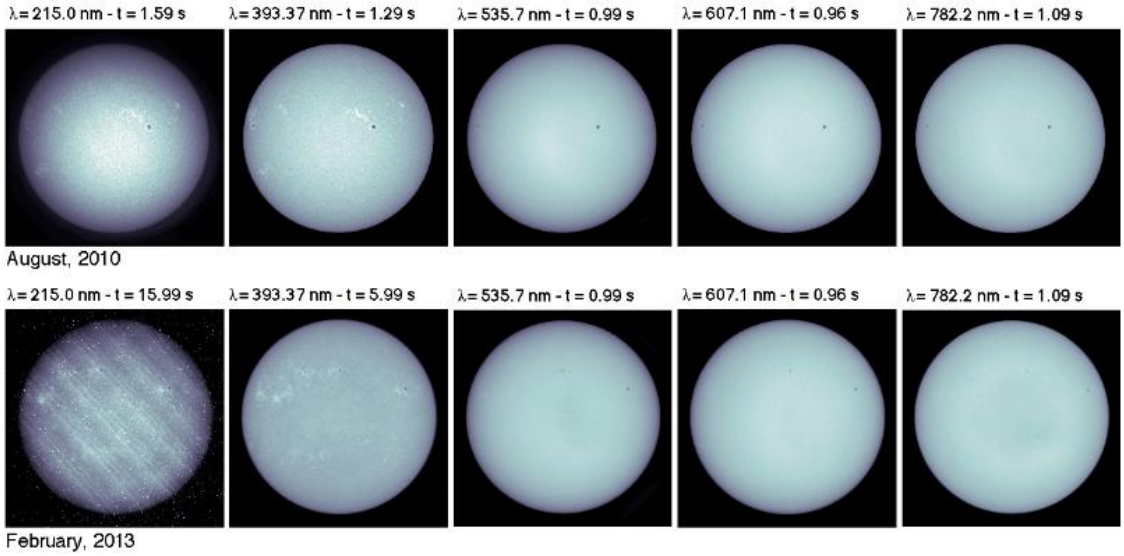


Figure 40: Sun's surface at different wavelengths and for various times in level 1 data products (Meftah et al., 2013).

The combination of solar irradiation and instrumental contamination badly affects the SODISM image quality, causing degradation. This is due to the harsh environment for optics, where several types of physical interaction lead to the potential for the severe degradation of optical performance over time (Meftah et al., 2015).

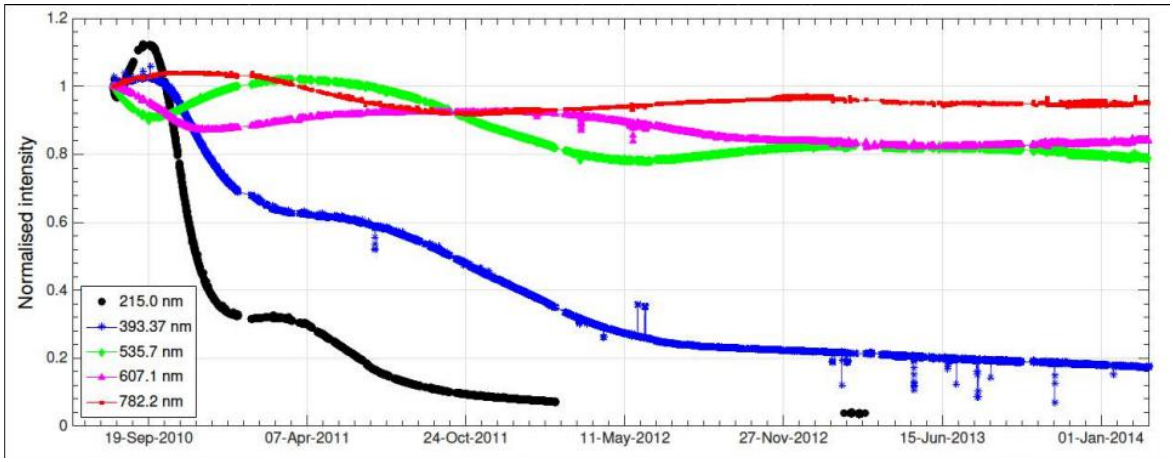


Figure 41: Normalized time series of integrated intensity of PICARD during its mission showing considerable degradation in the UV spectral bands (Meftah et al., 2015)

Channel 215 lost more than 90% of the normalized intensity, and the 393 nm channel lost about 80% (Meftah et al., 2013). Furthermore, pronounced degradation in the UV channels is shown in Figure 41. The reasons for the degradation could include the polymerisation of dust or contaminants on the front window, or on the other optical elements under exposure to solar UV; meanwhile, the near infrared and visible channels provide temporal oscillation but remain relative (Meftah et al., 2013).

About100 JPG format images were obtained at wavelength 215 nm, which were sized at 2048×2048 pixels. As previously mentioned, the images on this wavelength suffered from degradation and darkness, and it was explained that the application of methods to enable clear sunspot map images achieved limited success in this study, and thus have been postponed for future work.

3.7.1.2 Wavelength 393 nm

Since wavelength 393 nm is not suitable to measure the number of sunspots, the images recorded at wavelength 393 nm are used to determine the active regions near the solar limb that may alter the diameter measurements. Images with the Ca II line of 393.37 nm are used to

determine the active regions near the solar limb. Furthermore, these images are used to measure the solar differential rotation as well as space weather (SW) (Meftah et al., 2015).

Regarding the lifetime of the PICARD satellite, Figure 41 illustrates that wavelength 393 nm images were of better quality than those at 215 nm. However, they still suffered from degradation towards the end of PICARD's mission. In this work, a number of pre-processing steps were utilised to achieve satisfactory results. The images were in JPG format and their dimensions were 2048×2048 pixels. Approximately 300 images were downloaded. These images have different levels of contrast, and were also captured between August 2010 and January 2014. Figure 42 shows examples of the 393 nm images.

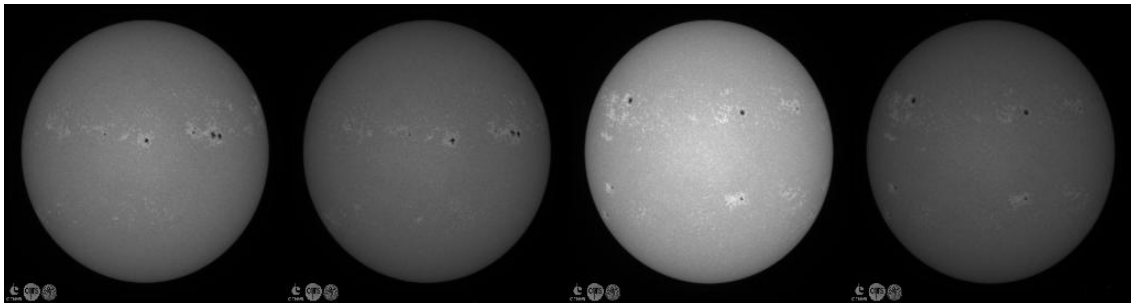


Figure 42: Examples of wavelength 393nm images showing contrast levels

The original 393 nm images have a lower contrast and less high-frequency detail, and the 8×8 blocking artefacts are far more visible. Distributed over the whole image are small bright and dark small spots of about one or two pixels in size; however, only the bright spots are visible outside the solar disc against the background, which is close to zero intensity. Most of these features are oriented at an angle of about -45° to the vertical, although a few were closer to the horizontal. Both dark sunspots and bright solar features are clearly visible on the solar disc.

The PICARD wavelength 393 nm images have been successfully corrected for limb darkness. Along with a horizontal line through the image centre, the background contrast varies from the centre of the solar disc to the limb near the edges. All the images downloaded have been measured in terms of their average radius and (x, y) centre. A 3×3 median filter was

applied to the original image and an ellipse was fitted to the solar disc of the filtered image, with its average centre at 1.04×10^3, 1.013×10^3, and the ellipse semi-axes at 1.039×10^3 and 1.012×10^3. The solar disc was re-centred at 1.0245×10^3, 1.0245×10^3 as a circle, with a radius at 0.92×10^3, and the limb darkening was removed through renormalizing, using the median of the radial values. This was the average for all images, but for accuracy, the information in the header FITS images have been used in the segmentation method for each. The location of the sunspots can be determined via spherical trigonometry; this depends on calculating the centre and radius of the solar disc. Furthermore, the radius and centre of the solar disc can be used for resizing and centring images in correction matrix applications. The result was an image (see Figure 43) with a flat solar background and with no linear background banding apparent.

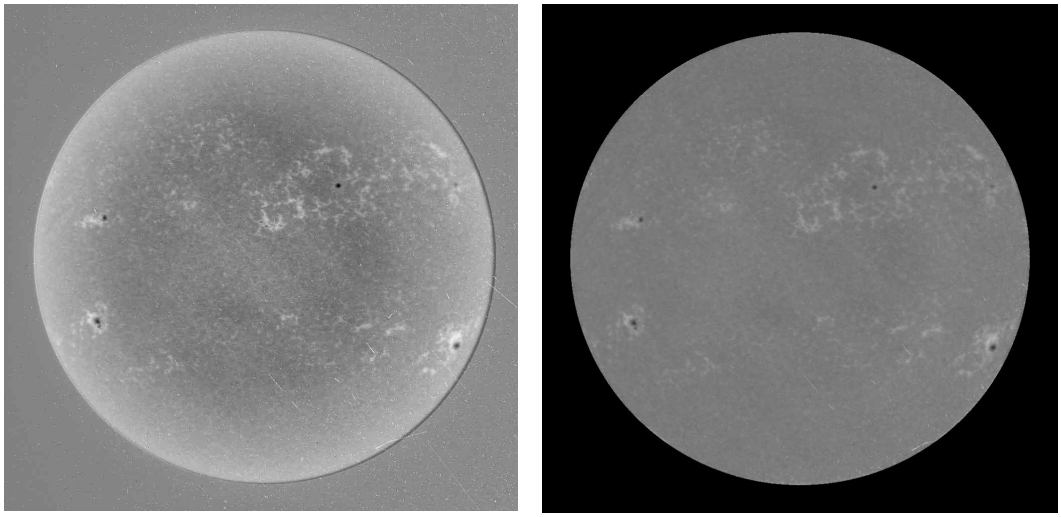


Figure 43: PICARD corrected 393 nm image on the left and in-house corrected image on the right

3.7.1.3 Pre-processing of sunspots detected at wavelength 393 nm

The pre-processing steps used included the dark limb removal. In fact, the behaviour of the light could have been affected by external factors. Indeed, physically, the manner of the light curve between the time the source centre enters the caustic, and the time the source edge enters the caustic is determined by limb-darkening parameters. The main factors that interfere with the shape of the light curves are as follows (Rahvar, 2015):

- The source star and its size.
- The components of the lens, mass and location.

- The Earth's location and its motion around the sun.
- The path of the source relative to the lens system.

The fundamental segmentation process at wavelength 393 nm shows in Figure 44, more details will provided in algorithm 1.

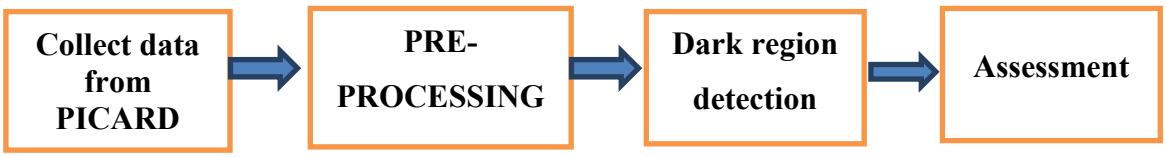


Figure 44: Segmentation images at wavelength 393 nm

The first step involves the application of the proposed enhancement method to remove unwanted noise and to provide a benchmark image. The method used in this segmentation applies two filters, namely the Haar and the Kuwahara filter. These filters have been adopted because, with suitable coefficients, the forms enable the further removal of most of the noise, while the structures are left unaffected with their details preserved. The Kuwahara filter[14] is a non-linear smoothing filter with good noise-reduction, which, at the same time, preserves the edges. A filter that accomplishes this aim is termed an edge-preserving filter. This results in smoothing the image while preserving the edges (Young, Gerbrands and van Vliet, 2007). Since edges are important when applying segmentation, the goals of the Kuwahara Filter are edge preservation and, at the same time, the capacity to apply smoothing on the image while preserving the edges, for example to understand basic of this filter assuming grey scale image *I(x,y)* and a square window of size *2a+1* centred around a point *(x,y)* in the image. In this case, the square will be divided into four smaller square regions *Q1, Q2, Q3,* and *Q4* and will be represented as:

[14] https://imagej.nih.gov/ij/plugins/kuwahara.html

$Q_1(x,y) = [x,x+a]\times[y,y+a]$

$Q_2(x,y) = [x,x-a]\times[y,y+a]$ (11)

$Q_1(x,y) = [x-a,x]\times[y-a,y]$

$Q_1(x,y) = [x,,x+a]\times[y-a,y]$

where X is the Cartesian product. In fact, any pixels located between two regions at the borders must belong to those regions, making a slight overlap between sub-regions.

The value of the central pixel will be determined from calculating the standard deviation $\sigma_i(x,y)$ and arithmetic mean $m_i(x,y)$ of 4 areas centred around pixel (x,y). The window is partitioned into four regions, as shown in Figure 45. The results after applying the Kuwahara filter $\Phi(x,y)$ for the point (x,y) is given by:

$$\Phi(\mathrm{x},\mathrm{y}) = \begin{cases} m_1(x,y)\ if\ \sigma_1(x,y) = \sigma_i m_i(x,y) \\ m_2(x,y)\ if\ \sigma_2(x,y) = \sigma_i m_i(x,y) \\ m_3(x,y)\ if\ \sigma_3(x,y) = \sigma_i m_i(x,y) \\ m_4(x,y)\ if\ \sigma_4(x,y) = \sigma_i m_i(x,y) \end{cases} \quad (12)$$

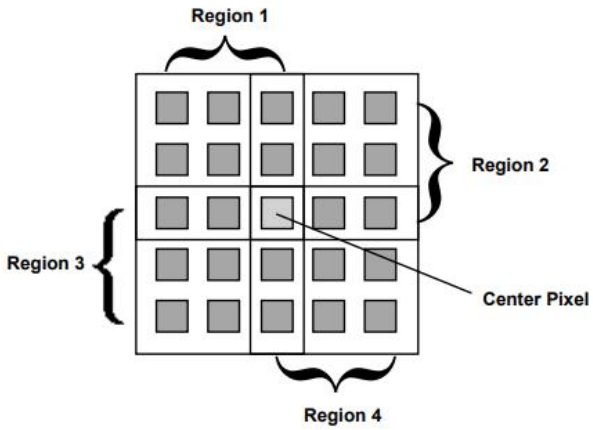
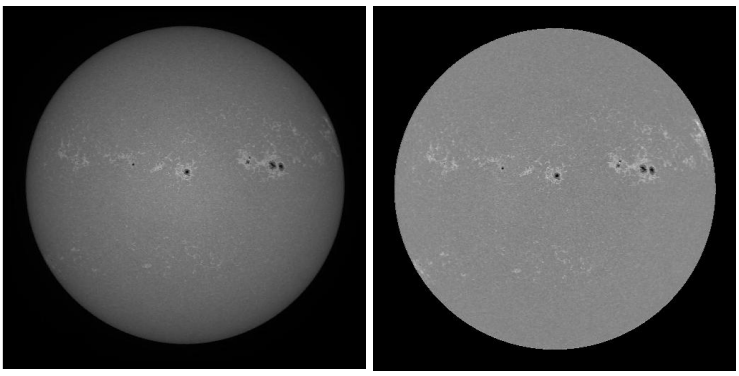


Figure 45: Four square regions defined for the Kuwahara filter

The asymmetric square neighbourhood around every pixel of a grey level image is divided into 4 square sub-regions, and the new value of the central pixel will be the grey level average over the most homogeneous sub-region. Particularly, this makes the central of the pixel have the mean value of the highest homogenous (Young, Gerbrands and van Vliet, 2007). the position of the pixel in relation to an edge will determine which region will have a greater standard deviation, for instance, the mean value of the dark

region will probably be taken if the pixel is located in a dark side, on the other hand, the opposite should be true for a bright side (lighter side of an edge it will most probably take a light value). Moreover, if the pixel is located on the edge it will take the value of the smoother, least textured region (Kuriakose and Joy, 2014). Figure 45 shows the principle of the Kuwahara filter and figure 46 shows the results after using this method.



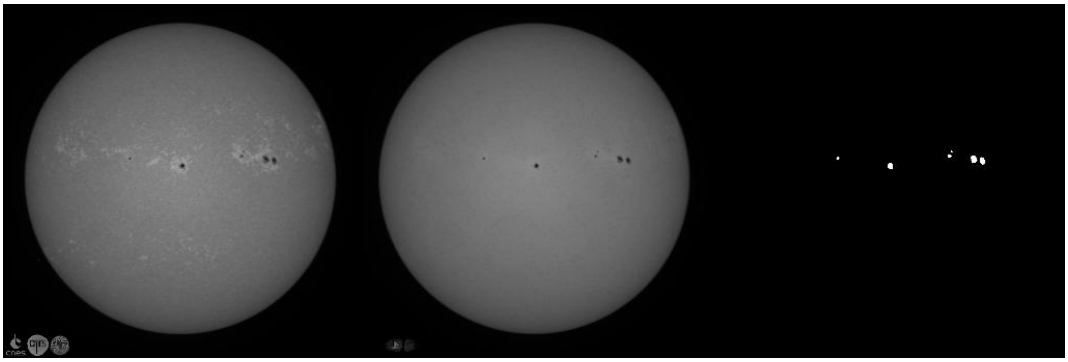
(a) (b)

Figure 46: (a) Example of original image at wavelength 393 nm; (b) After limb darkening removal

The procedure applied to JPG images at 393 nm can be summarized as follows for algorithm 1:

Algorithm 1: Sunspots detection of 393 nm

i. Both the Haar wavelet filter and Kuwahara filter have been applied to the enhanced images.
ii. Outliers are removed with a threshold of 2 to remove bright regions.
iii. The band-pass filter and Gaussian filters are applied.
iv. Conversion to binary image is conducted. Figure 47 shows an example of the results achieved using this segmentation procedure.



(a) (b) (c)

Figure 47: Example of segmentation at wavelength 393 nm (a) original image; (b) After removal of outliers and (c) Final results

Figure 48 shows the flowchart of this method applied to wavelength 393 nm data to detect sunspots.

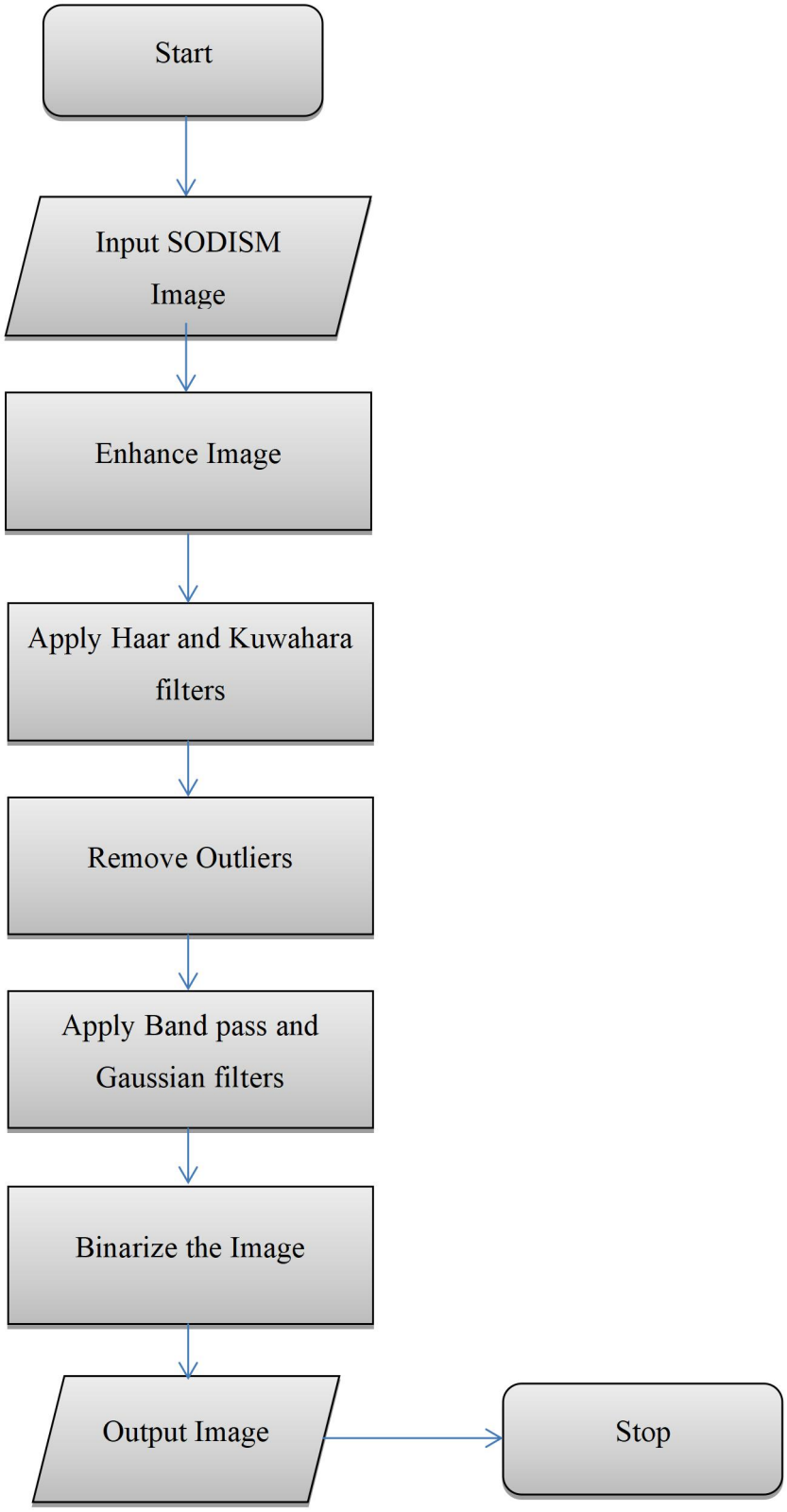


Figure 48 Flowchart for sunspot detection at 393 nm

3.7.1.4 Wavelength 535 nm

Approximately 203 images in jpg format were obtained at wavelength 535 nm; these were sized at 2048×2048 pixels and collected during the period August 2010 to January 2014. This wavelength is located in the visible spectral band. Figure 41 shows the normalized time series of the integrated intensity of PICARD during this mission, and illustrates degradation at the end of mission. These images appear to have greater clarity and definition compared with previous wavelengths because this pass-band is quasi-free of Fraunhofer lines (Meftah et al. 2014). Figure 50 shows examples of 535 nm images.

The original 535 nm image has a lower contrast. After performing limb-fitting on the original image, it was found that the average solar centre is 1.04×10^3, 1.012×10^3, and it was re-centred to 1.024×10^3 and 1.024×10^3, with a radius of 0.92×10^3. These values have been obtained from many of the images tested. As previously mentioned, for more accuracy, the x and y values are chosen from the header in the FITS image. The normalised version exhibits some radial contouring which indicates the need to use a function fitted to the radial estimate of limb darkening to perform the correction. The 535 nm images are located in a visible wavelength and the solar emissions are less intense towards the edge of the sun in this wavelength. The intensity variations distribute the intensity across the disc to avoid dark limbs, Although limb darkening removal is required for filament detection, there is no problem in detecting the active regions, such as sunspots (Qahwaji and Colak, 2005a). The result obscures some of the background features that are visible in the PICARD corrected image. The PICARD corrected image (see Figures 49 and 50) shows sunspots and the surrounding faculae, but also some non-radial background variation and slight banding in the background at an angle of about -30° to the vertical in the lower-left half of the image. There is a dark ring around the edge of the solar disc, and the area outside the solar disc is grey with brighter patches in all four corners. These background variations may make it difficult to correctly segment some faculae. There are many dark spots visible on the solar disc (not as dark as sunspot umbra and a few pixels in size) and a dark

linear feature of about 48 pixels long is located near the bottom-right edge of the disc.

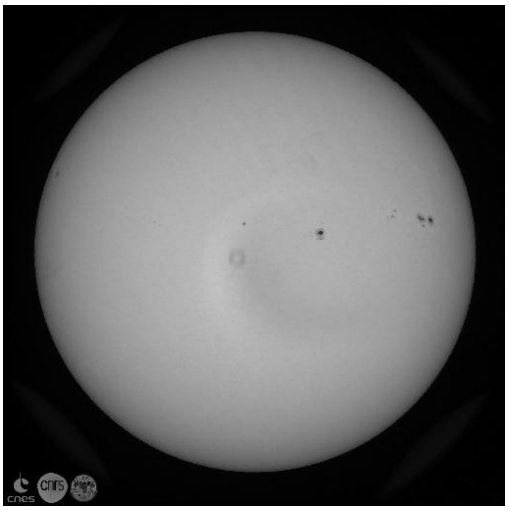


Figure 49: Example of an image at wavelength 535 nm from 2nd October 2011

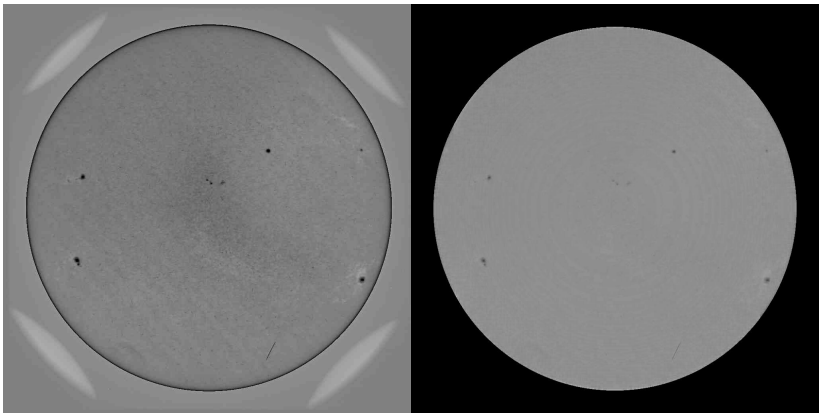


Figure 50: PICARD corrected 535 nm image on left and in-house corrected image on right

Despite the fact that the 535 nm images are generally better visually than those at the previous wavelengths, some distortions still exist; for example, in the ghost effect and lines appearing on the images illustrated in Figure 51, which makes the task of segmentation challenging.

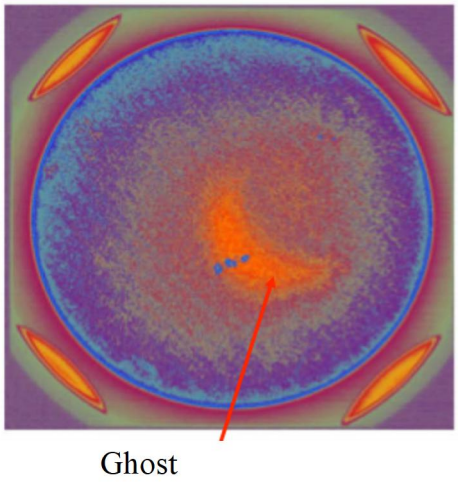


Figure 51: Ghost in an image at wavelength 535 nm

The approach adopted to detect sunspots is summarized in Algorithm 2.

Algorithm 2: Sunspot detection at 535 nm	
i.	Detect the solar disc and record its centre and radius information.
ii.	Eliminate all information located outside the solar disc.
iii.	Convert the image scale from signed 32 bit to unsigned 8 bit; this helps to reduce computational complexity to make them easy to handle.
iv.	Remove noise and some other unwanted features from images by applying a combination of wavelet-based Kuwahara and À Trous filters.
v.	Apply filtering to remove outliers.
vi.	Apply a band-pass filter to display the dark regions on a normalized background.
vii.	Identify the threshold to obtain a mask image that determines the sunspot locations.
viii.	Detect the sunspots by binarising the image.

An example of the results from this method is shown in Figure 52.

Figure 52: Sunspot detection for an example of an image at wavelength 535 nm

To visually compare the segmentation performance, a comparison between the original and segmented image is shown in Figure 53. Furthermore, Figure 54 shows a sample image with a ghost effect and results; even though there is a ghost effect, the method still gives acceptable results.

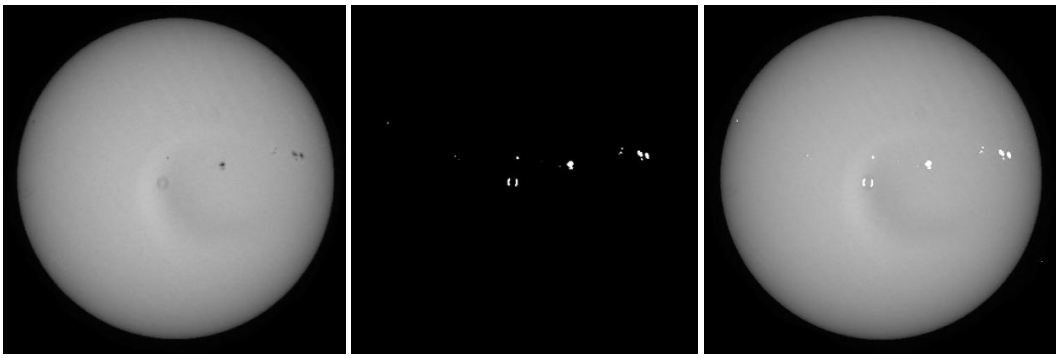


Figure 53: Original image (left) segmented image (middle), showing sunspots positions on original image for wavelength 535 nm (right)

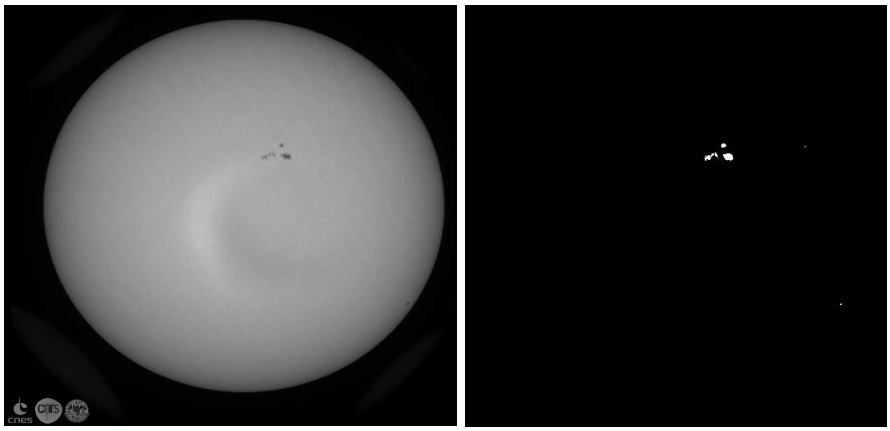


Figure 54: SODISM 535 nm image from 12:30, 1st October 2011 (left) and resulting detected sunspots (right)

3.7.1.5 Wavelength 607 nm

Approximately 130 images in JPG format were obtained at a 607 nm wavelength; these were sized at 2048×2048 pixels and retrieved between August 2010 and January 2014. The original images have a lower contrast and less high-frequency detail, probably due to degradation. All the PICARD JPEGs contain three components and two quantization tables for the intensity and colour components. The values in these tables range from 18 to 215 and 30 to 176 for image 215; from 20 to 215 and 30 to 176 for image 393; from 20 to 215 and 30 to 176 for image 535; and from 6 to 70 and 10 to 57 for image 607. These JPG quantisation tables are designed to remove image data that are less visible to the eye and unhelpful if the images are to be processed automatically. This band is similar to the 535 nm wavelength in that it is located in the visible wavelengths, and therefore suffers from less intensity towards the edge of the Sun (Figure 54). The corrected versions will attenuate the intensity across the entire disc of the Sun. Hundreds of images that test the performance of a limb fit to the original image find the solar centre average at 1.037×10^3, 1.011×10^3. The solar disc was re-centred at 1.024×10^3, 1.024×10^3 as a circle with a radius of 0.915×10^3; the limb darkening was removed through renormalizing, which was achieved by using the median of the radial values. After suitable contrast stretching (output 0 to 255 for input 56 to 224) the normalized image is ready for processing. It would be possible to work with these images, either with the original or the limb darkening corrected versions. However, the corrected version in Figure 56 shows some residual non-radial background variations; these could make it more difficult to extract some features. Nevertheless, sunspots and some faculae should be extractable. The images at this wavelength are of a high visual quality because the features are clear and the noise seems hidden. Indeed, SODISM bands 393 nm,535 nm,607 nm and 782 nm are of sufficient quality to feed the solar spectral irradiance (SSI) models (Meftah et al. 2016) and there is no ghost effect. They still have the brightest centre, although the edges and noise need enhancement before the method to retrieve the sunspot maps can be applied. Figure 55 shows an example of the 607 nm images.

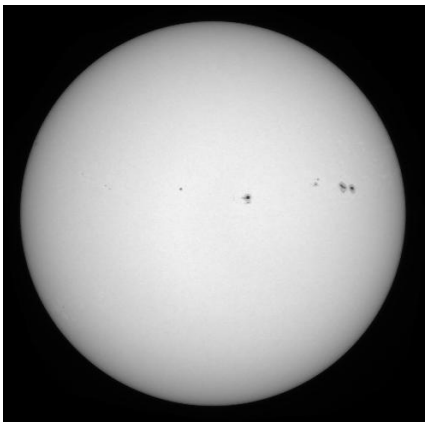


Figure 55: Solar disc as an example of an image at wavelength 607 nm at 1st October 2011

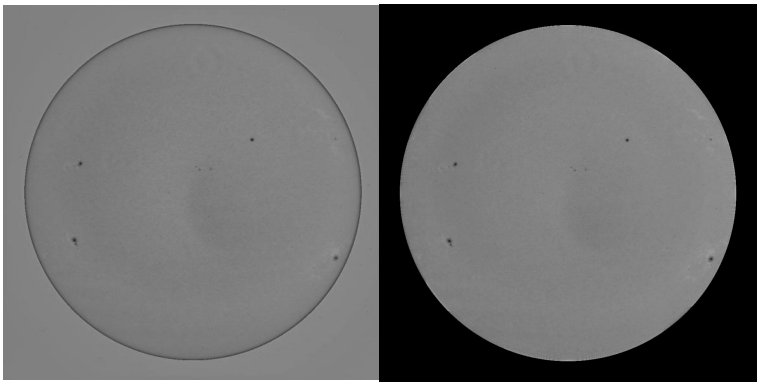


Figure 56: PICARD corrected 607nm image on left and in-house corrected image on right

Algorithm 2 is also used to achieve segmentation, and the sample segmentation results are shown in Figure 57.

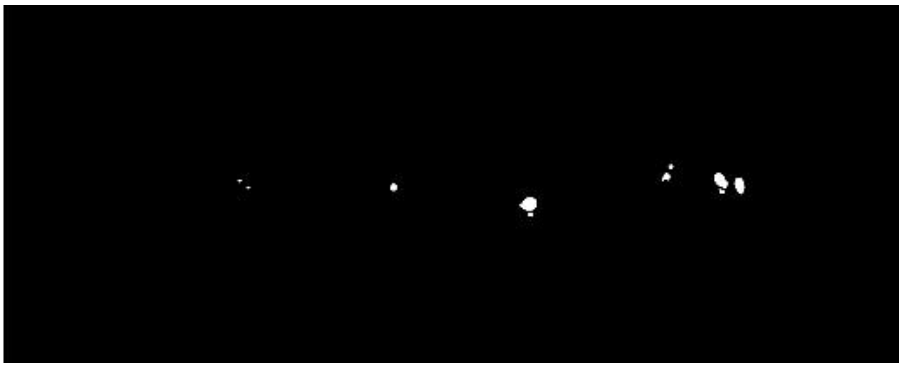


Figure 57: Sunspot detection for an example of an image, wavelength of 607 nm

Figure 58 shows the sunspot detection matched with the original image; this is a typical example of 607 nm.

Figure 58: Original image (left) sunspots detected (middle) and sunspots positions on original image (right) for an example of an image at wavelength 607 nm

It was observed that the algorithm-detected sunspots were very few and relatively small in size (Figures 59 and 60).The visual inspection for random samples of SODISM images was consistent with the previous results, which means that this method has achieved its aims.

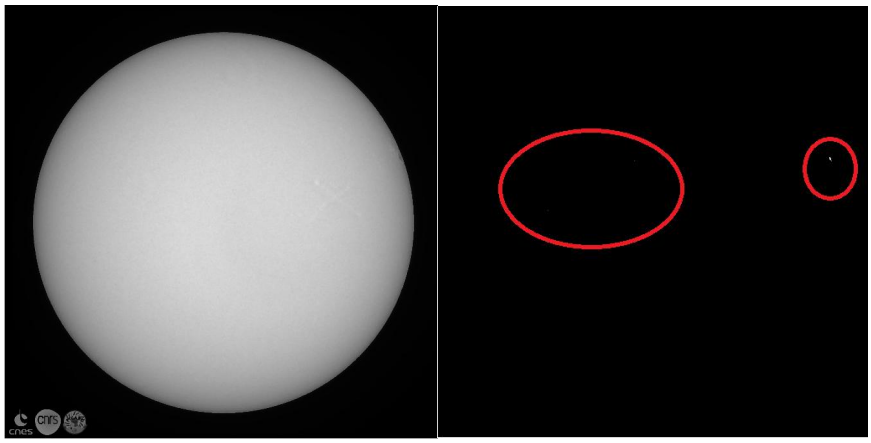


Figure 59: Input image at wavelength 607 nm from 4th October 2010 06:41 (left) and sunspot detection image (right)

Figure 60: Magnified sunspot detection in an example of an image at wavelength 607 nm from 4th October 2010

3.7.1.6 Wavelength 782 nm

In total, 125 images were obtained at a wavelength of 782 nm. They were the same size as the images from the other wavelengths. The images at this wavelength are free from visible noise, and ghosting artifacts were not observed. Figure 61 shows an example of a 782 nm image.

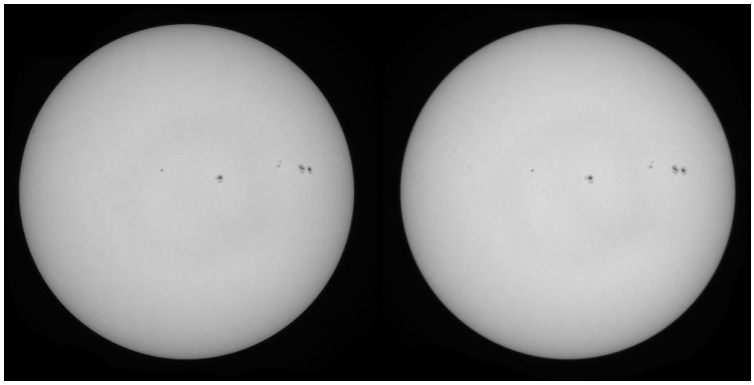


Figure 61: Solar disc as an example of an image at wavelength 782 nm from 1st October 2011

Algorithm 2 was used again to achieve the segmentation, and the sample results are shown in Figure 62.

Figure 62: Sunspot positions on original SODISM image at wavelength 782 nm from 1st October 2011

3.7.2 Segmentation of the FITS images

In 2014, the website used to acquire PICARD images was updated, and this time all JPG images were removed. Only images at 535 nm in the FITS format were uploaded. Interestingly, the data retrieved was of a better quality and included meta-information that was contained in a header file, such as the radius and the x and y centres of the image. This method was applied to images of 535 nm, 607nm and 782 nm wavelengths, which were available at

level 1B[15]. Level 1B data products include a number of corrections for any instrumental issues. A total of 206 images were downloaded, and each image had a size of 2048×2048 pixels. The data obtained was from 5th August 2010 to 4th January 2014. After contacting the PICARD team, the author was provided with 306 images taken at wavelengths of W.L 607 nm, and 300 images at 782 nm; these were collected between 2010 and 2014.

Detecting sunspots as individual elements are equivalent to determining which pixels belong to each sunspot, thus determining which pixels are active points by using the accurate threshold intensity level to isolate these points from the background (Curto, Blanca and Martínez, 2008). The threshold should be computed for each image because it varies from image to image.

The method detects solar disc edges then it detects sunspots candidates and by iterative threshold on the gradient image which has been previously normalized counting the number of the connected regions, and figure 63 shows the flowchart diagram for the basic fundamentals of the applied segmentation method to detect sunspots.

The automated detection of solar features is a technique applied to provide robust, fast and accurate automated detection (Qahwaji and Colak 2005b). Many researchers have used observatory images from satellites, including SOHO and SDO, to detect features such as sunspots. However, since images from SODISM have been relatively underused, this study has worked with them. The method developed to automatically detect sunspots in SODISM L1 images at 535.7 nm and 607 nm is detailed in the section below.

[15] http://PICARD.projet.latmos.ipsl.fr/files.php

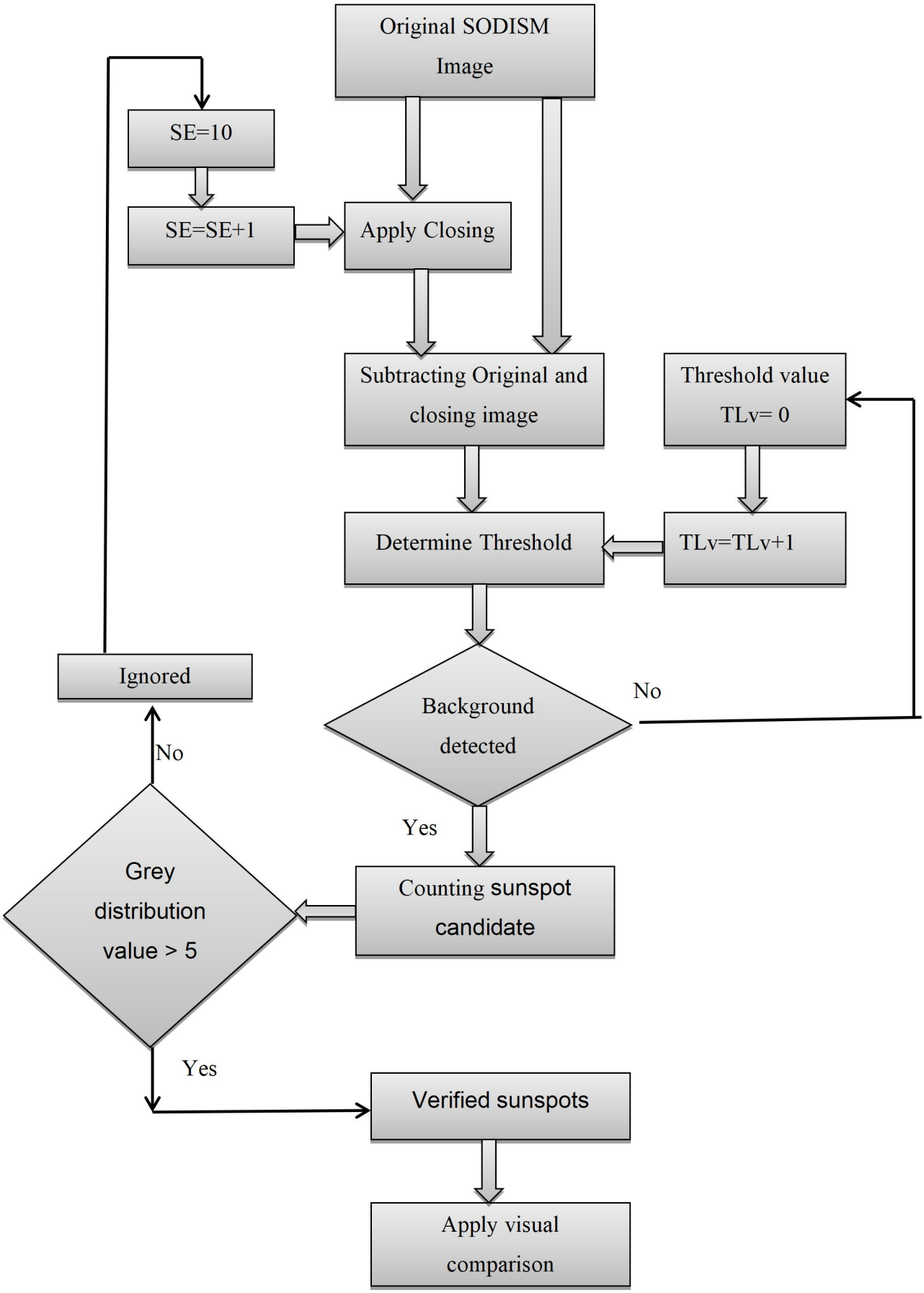


Figure 63: Chart diagram of the sunspot detection procedure.

3.7.2.1 Extraction of the solar limb in the FITS images

Morphological methods and the Otsu algorithm have been employed to calculate sunspot locations and their areas. To achieve this, the solar centre and radius have been calculated by extracting the solar limb. The steps involved are detailed in Algorithm 3.

Algorithm 3: Extraction of solar limb in Fits images

i. Achieve a clean disc without noise and sunspots by applying a closing operation with a structuring element (SE) that is applied to an original SODISM image.

ii. Choose a circular SE with a radius of 30 pixels (this value was chosen using cross-validation, and because the radius of the largest selected sunspot for all images is calculated as 30 pixels); the sample result is shown in Figure 64b.

iii. Determine the border edges to identify solar limb; from this, the first step shrinks the solar disc by one pixel (see the filtered image in Figure 64b) to produce an image smaller by one pixel. Subtract the new image from the filtered image, (the result is illustrated in Figure 64c).

iv. Eliminate CCD noise by utilizing the Kuwahara filter (see Figures 64d and 64e).

v. Use the binary overlay plugin between the original image and the solar limb image which is labelled in red and overlaps the original image, as shown in Figure 64f.

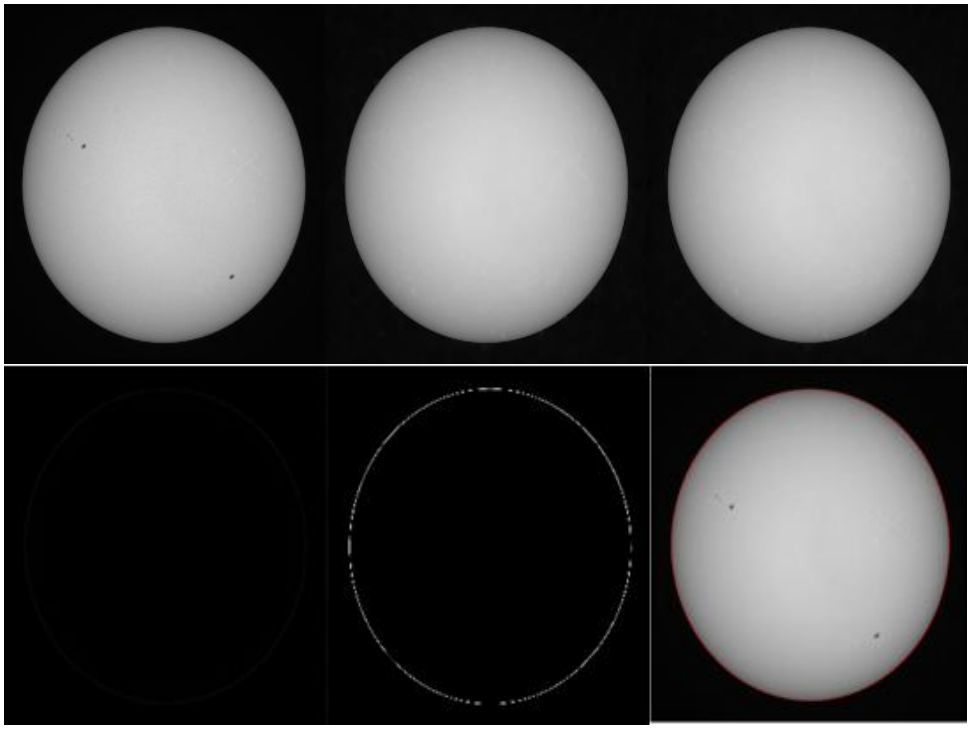


Figure 64:Benchmark image: (a) Filtered image; (b) Shrunken image of the solar disc; (c) With a radius 1 pixel smaller than that in (b); (d) Solar limb shown in grey image; (e) Solar limb shown in a binary image; (f) Red around the disc shows the solar limb label

.

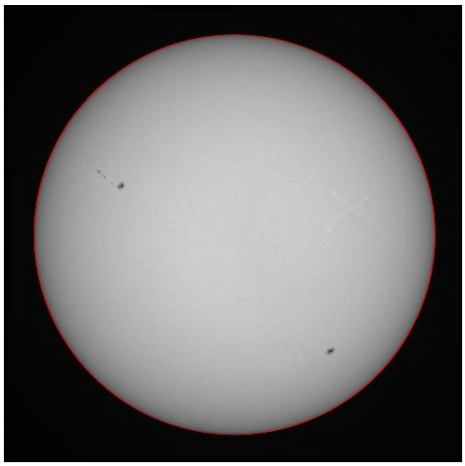


Figure 65: example of image at wavelength of 607 nm at 25th September 2010 time 00:07, with the solar limb labelled in red and overlapped on Figure 64(a)

3.7.3 Detection and verification of sunspots

Sunspots detection and verification involves the recognition of sunspots in the solar disc after the solar limb has been extracted. Due to the limited resolution of the data, the sunspot umbra and penumbra are not separated in

the SODISM images; rather, they are considered and processed as an integral part of the sunspot. The steps outlined in Algorithm 4 (see Figure 66) enable the recognition of sunspots.

Algorithm 4: Detection of sunspots

i. Process the original SODISM image using the proposed quality enhancement method; an output sample is shown in Figure 66(a).

ii. Compute the gradient of the boundaries of sunspots (see Figure 66(a)).

   a. Fill holes with a closing operation; this leads to the removal of dark regions surrounded by bright crests in grey scale images.

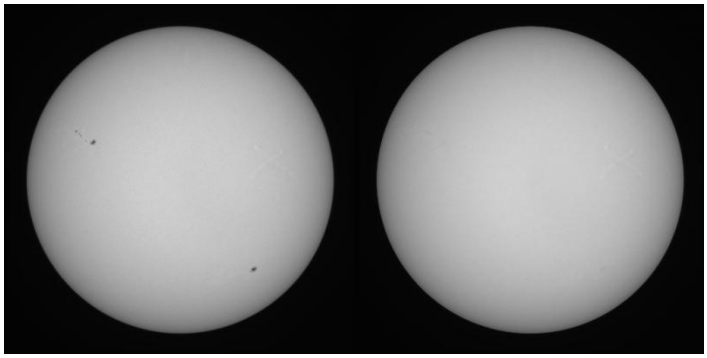
   b. Compute the image obtained in a. and the gradient image obtained in [ii], which is the difference between Figures 66(a) and 66(b) and yields 66(c).

iii. Separate the sunspot gradient from the noise, as shown in Figure 66(c). This provides the darkness threshold, shown in Figure 66(c). Experiments have been conducted to determine a suitable value, and 15% intensity was selected, as shown in Figure 65. Due to the solar limb darkening, the sunspot gradient is lower at the solar limb, and therefore the threshold will be 10% in the solar disc region.

iv. Remove unwanted noise using the Kuwahara filter and employ the extended min and max operation for marker detection following segmentation. Figure 66(d) shows a sunspot candidate.

v. Acquire sunspots from candidates in Figure 66(d). The candidates are considered verified sunspots if the difference between the maximum and the minimum grey distribution value of a pixel is greater than 5 and the other regions are ignored.

vi. To make a visual comparison, apply a binary overlay in red, and superimpose the original image. However, this is not a compulsory step. (Figure 67 shows sample result)



(a) (b)

(c) (d)

Figure 66: (a) Original image disturbed by instrument noise; (b) Clean image without sunspots(c) Gradient on the image; (d) binary image showing sunspot candidates

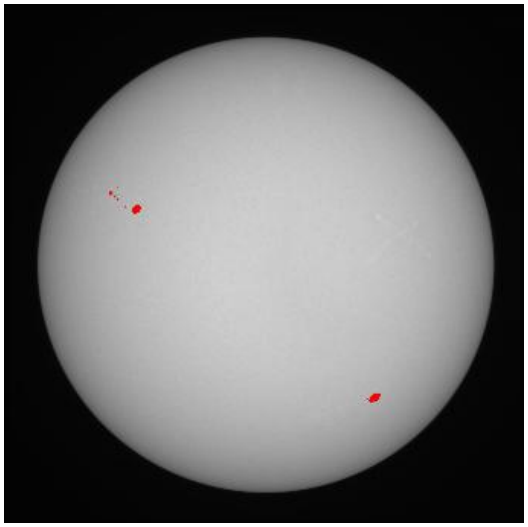


Figure 67: Recognised sunspots superimposed on the original image at 535 nm from 4th January 2010 at 21:21.

A visual assessment was made to evaluate the final results, as illustrated in Figures 66 and 67. This is applied to random images (approximately 100) using the aforementioned method, even though the assessment shows matches, as shown in the previous figures. However, the method needs further verification. The verification method deployed

compares the results of the proposed automatic detection of sunspots and those in a NOAA catalogue[16], which is based on existing data. Examples of the input solar image and the corresponding detected sunspots are shown in Figure 68. An example of a resulting image at a 535 nm wavelength is shown in Figure 68, where the original image was captured on 4th January 2010 at 21:21.

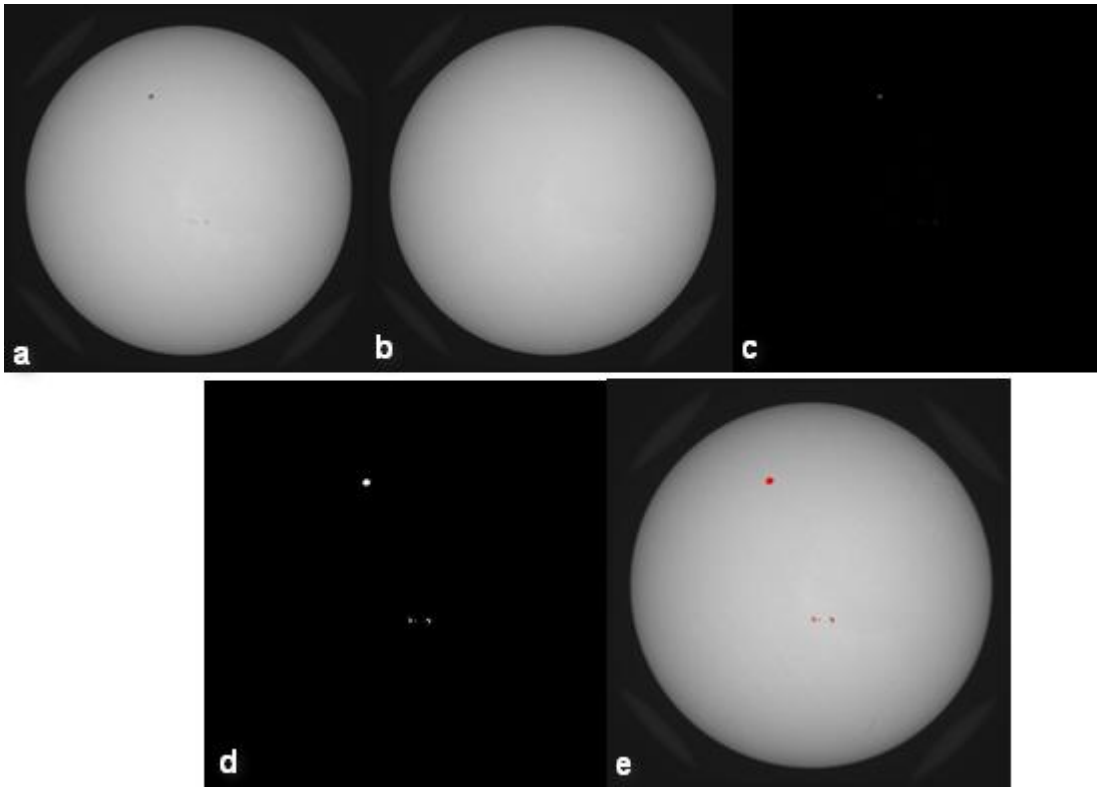


Figure 68: Image at wavelength 535 nm from 4th January 2010 at 21:21 (a) Original image disturbed by instrument noise; (b) clean image without sunspots; (c) gradient on the image; (d) binary image showing sunspots candidates; (e); recognised sunspots superimposed on the original image

The data at wavelength 535 nm from August to September 2010 for the automated methods are calculated and listed in Table 7. The first column shows the date the image was captured, and next to it is the number of sunspots detected by the automated method; the third column shows the time at which the SODISM image was taken. The fourth column shows the number of sunspots counted by the NOAA, and the fifth column shows a suitable time chosen in the NOAA catalogue. The sixth column shows the false rejection rate (FRR), which is the number of sunspots detected by the

[16] https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/sunspot-regions/usaf_mwl/

automatic system but not by the NOAA inspection. Finally, the last column shows the false acceptance rate (FAR), which is the number of sunspots detected by NOAA but not by automatic detection. The rationale for computing the FRR and FAR is to evaluate the results of the proposed technique against reliable findings, such as those by the NOAA (Zharkova et al. 2005) The comparison included the sunspot's position which is shown in the third column in NOAA; the method, which was adapted to identify the sunspot's position in SODISM, depends on measurements of the radius and angles to determine the sunspot's exact position. Java script calculation of the solar orientation angles and solar latitudes and longitudes of sunspots[17] was employed to detect the results. Algorithm number 5 is described to find associations and make comparisons between sunspots in SODISM and NOAA catalogues as shown in the flow chart of figure 69. The algorithm 5 has been applied to obtain a comparison between the number of sunspots detected in the SODISM images and the NOAA catalogue.

Algorithm 5: Comparison between the number of sunspots

i. Find sunspot groups that match the NOAA date.
ii. If the time is six hours or less the data is recorded.
iii. If there is more than one classification record for the sunspot group time within the six hours, then the one with the minimum time difference is chosen.
iv. Calculate the sunspot's position in the chosen SODISM image and compare this with NOAA
v. Calculate the accuracy between the proposed method and the NOAA catalogue by the correlation coefficient.

[17] http://www.nature1st.net/bogan/astro/sun/sunspots.html

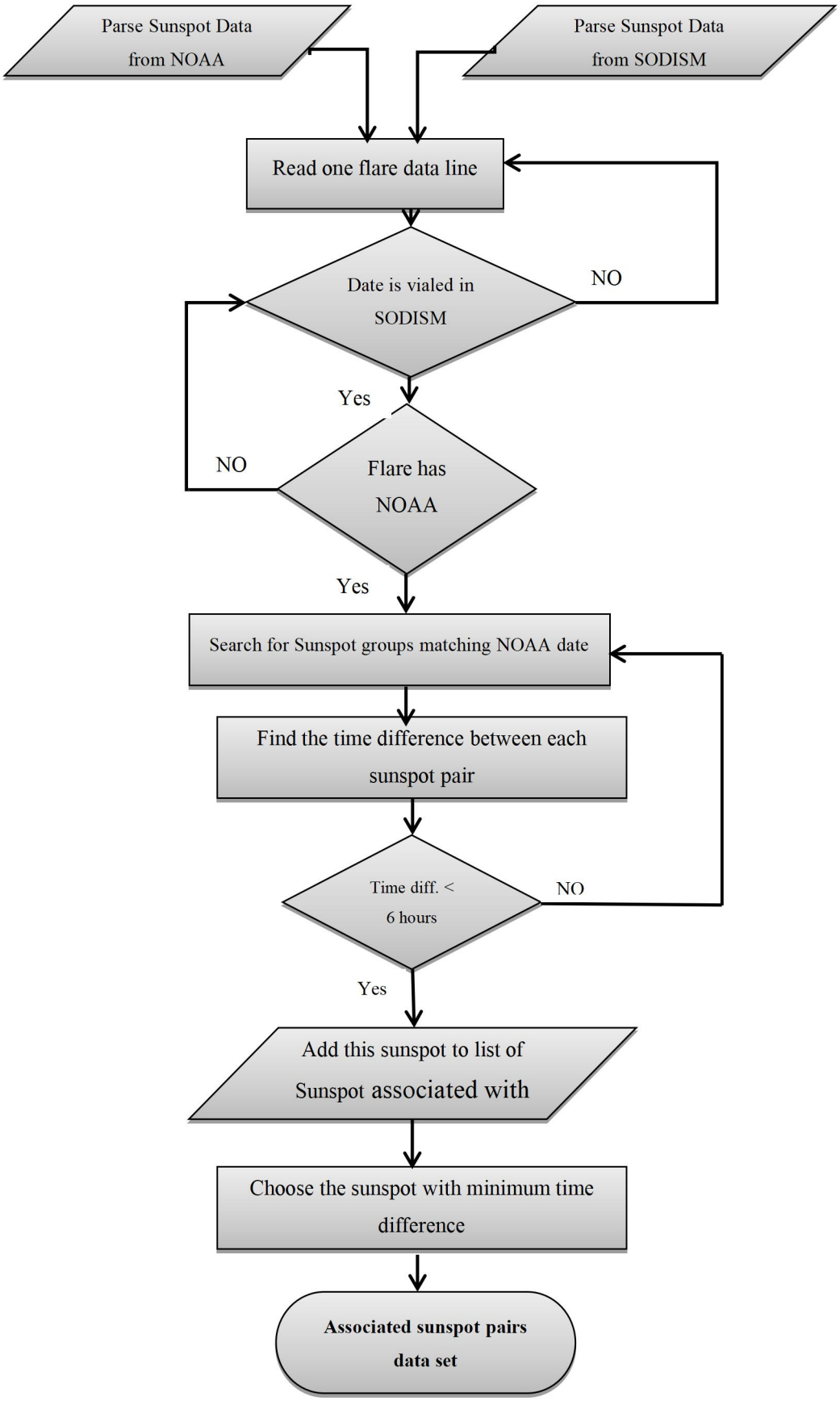


Figure 69: Comparison between numbers of sunspots

Depending on the NOAA catalogues, the accurate sunspot number results can be identified; Figure 70 illustrates the structure of this catalogue. The algorithms were applied to the sunspot data for the period from 2010 to 2014, and Table 7 shows an example for the August-September 2010 period.

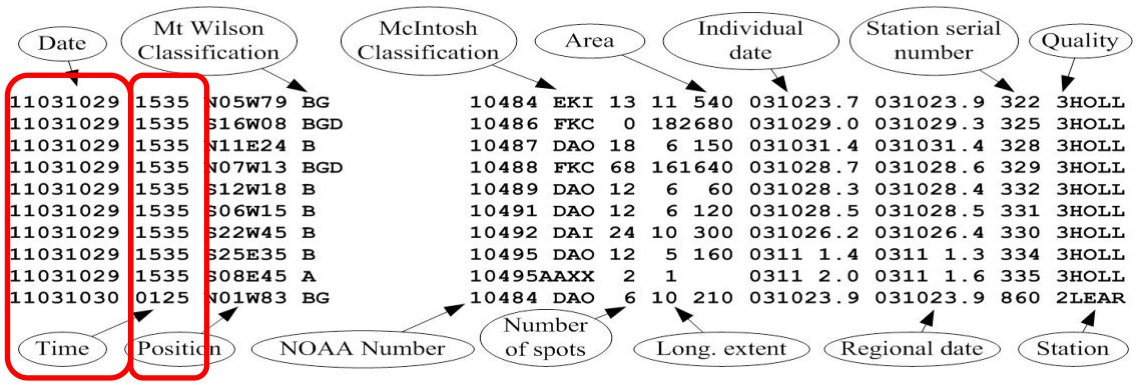


Figure 70: NOAA sunspot catalogue data

The example in Figure 71 shows an example of a sunspot recorded in a SODISM image on the 5th August 2010 at 04:49. Moreover, searches in the NOAA catalogue for 5th August 2010 show data for three different times, namely: 02:36, 07:35 and 16:35. Thus, the nearest time at 07:35 is chosen, the number of sunspots in the NOAA was four, and the SODISM segments map also show four; furthermore, the position of these sunspots is recorded as N13W24, N11E67, S19E56 and N25W49 respectively in the NOAA catalogue, which matches the results from the SODISM images inspection.

```
11100805 0236 N12W21 A          11092 HHX  1  3 200 100803.5 ------.- 062 2LEAR
11100805 0236 N10E68 B          11093 DAO  4 10 160 100810.2 ------.- 063 2LEAR
11100805 0236 N25W46 B           //// CSO  8  4  40 100801.5 ------.- 064 2LEAR
11100805 0735 N13W24 A          11092 HHX  1  3 170 100803.5 ------.- 054 3SVTO
11100805 0735 N11E67 B          11093 CAO  4  8 150 100810.4 ------.- 055 3SVTO
11100805 0735 S19E56 A           //// AXX  2  0  10 100809.6 ------.- 056 3SVTO
11100805 0735 N25W49 B           //// DSO  5  5  50 100801.5 ------.- 057 3SVTO
11100805 1635 N13W30 B          11092 CSO  2  4 230 100803.4 ------.- 063 4HOLL
11100805 1635 S18E50 B           //// BXO  2  1  10 100809.5 ------.- 065 4HOLL
11100805 1635 N12E62 B          11093 CSO  4  8 190 100810.4 ------.- 066 4HOLL
11100805 1635 N25W55 B          11094 CSO  6  4  60 100801.4 ------.- 067 4HOLL
```

Figure 71: Example, 5th August 2010, for NOAA catalogue shows time chosen

The concept of measuring a sunspot position is shown in Figure 72. The coordinate system is known as heliographic and is expressed as the latitude Θ and longitude Φ of the Sun's surface; this can be extended to a three dimensional measurement through the addition of a radial distance between the centre of the Sun and the sunspot. Although this system is based on the

established work of Carrinsion (1863), the coordinate system is well-suited for SODSIM images when detecting sunspot positions (Thompson, 2006).

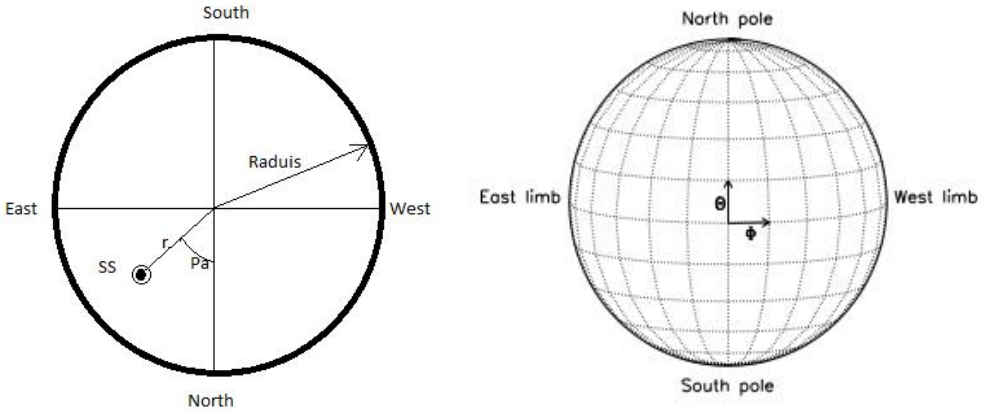


Figure 72: Measuring sunspot position (left), and heliographic longitude and latitude on the solar disk (right)

A random visual inspection of images confirms the results of SODISM catalogue.

3.8 Feature extraction and classification of sunspots

After sunspot segmentation, the number of attributes of each sunspot group is categorized according to the Modified Zurich classification system. This step is necessary when comparing the number of sunspots segmented using the proposed method with the data in the NOAA catalogue. A comparison between the numbers of sunspot groups with NOAA's findings is confirmation of the accuracy of the proposed method (p). In this research, two methods were used to compare the findings with the NOAA catalogue. The first method was to visually count the sunspot groups depending on the segmentation images, and the second method used the algorithm developed by Mehmmood et al. (2010), which is summarized in algorithm 6, and figure 73 shows the flowchart of extraction and classification of sunspot groups, while Figure 74 shows an example of the results.

The results as shown in figure 74, shows the classification of sunspots and gives accurate sunspot numbers. These results lead to the capacity to make accurate comparison with the NOAA catalogue, and to extract information such as correlation between SODISM and NOAA catalogues.

Algorithm 6: Extraction and classification of sunspot groups (Mehmmood A, at el. 2010).

The proposed segmentation method first masks the produced image so that the background is black and sunspots are white. The following steps represent the proposed segmentation method.

i. Compute the bounding rectangle that contains the sunspots.

ii. Check the size of each bounding box to find the one that contains adjacent sunspots.

iii. Employ an iterative algorithm to combine the group of sunspots.

iv. Adjust the maximum bounding box to contain the neighbouring boxes.

v. Repeat the above algorithm until it finds all spots contained in each group and fits them into one box. Then extract the area that contains the sunspot group from the main image, and save this as a separate image that represents one sunspot group.

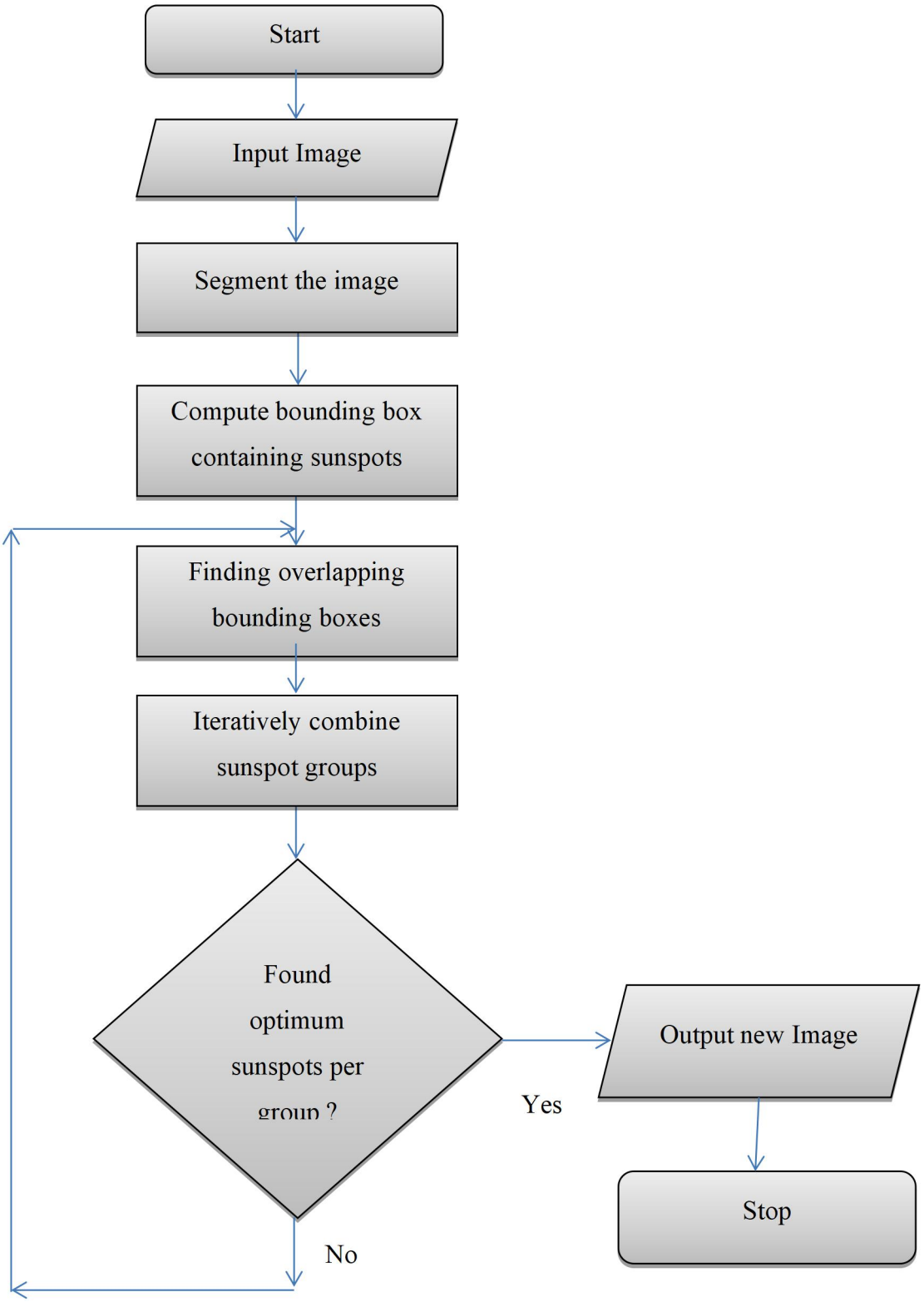


Figure 73: Flow chart extraction and classification of sunspot groups

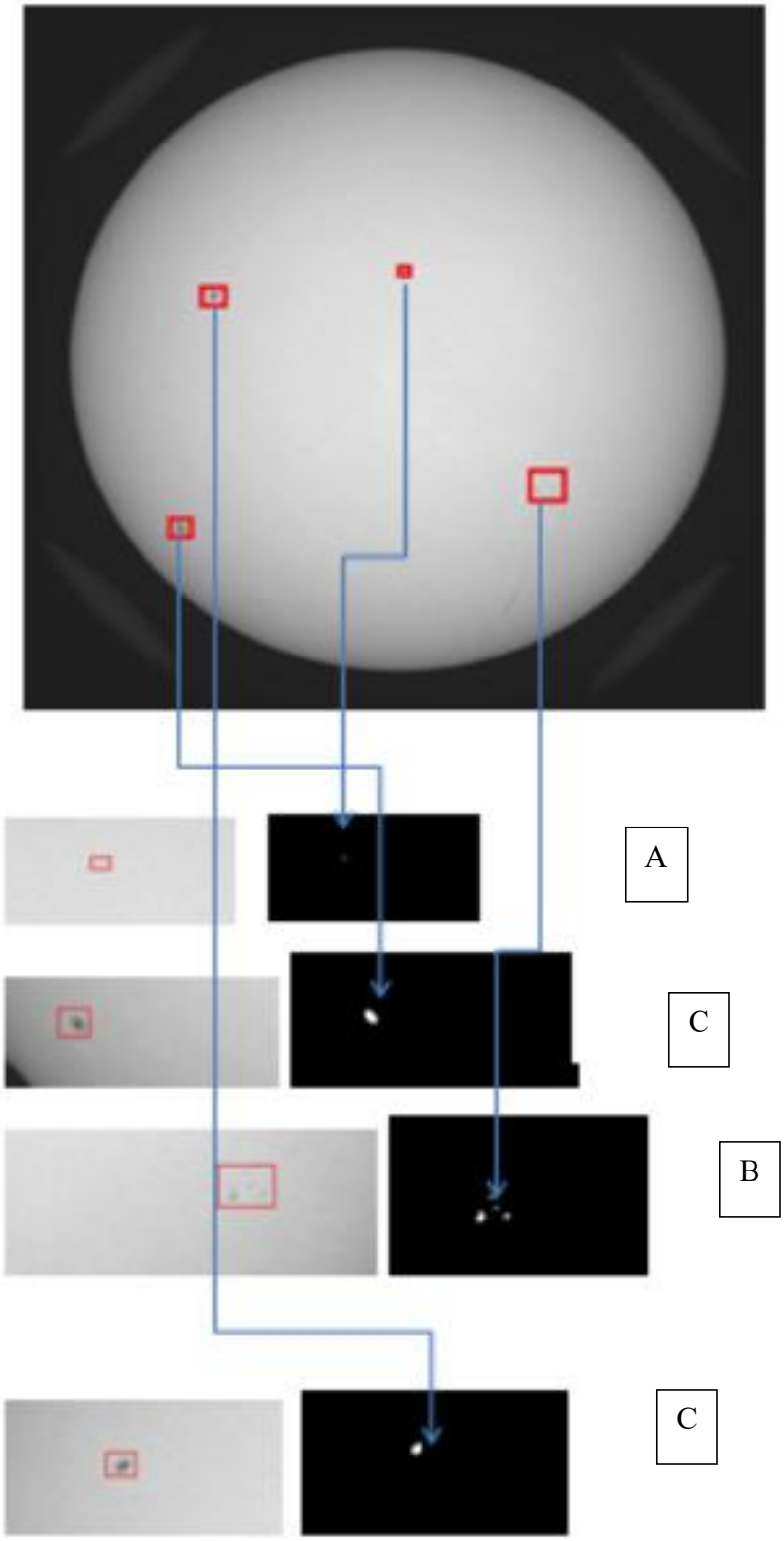


Figure 74: Sample showing 4 sunspots: the sunspot group classification technique applied at wavelength 535 nm according to MacIntosh on 17th October 2010 at 02:05, class A unipolar no penumbra, class B bi-polar no penumbra and class C bi-polar with penumbra

Table 7: Comparison of sunspots detected by the automatic procedure and the NOAA from August 2010 to September 2010

Date	Number of sunspots detected by automatic method	Time of Sunspot in SODISM image	Number of sunspots in NOAA catalogue	Time of NOAA Sunspot	FRR[18]	FAR
05/08/2010	4	04:49	4	07:35	0	0
06/08/2010	4	01:07	4	02:30	0	0
07/08/2010	4	01:07	4	02:37	0	0
09/08/2010	4	05:27	4	06:15	0	0
10/08/2010	4	01:17	3	02:00	1	0
11/08/2010	5	05:43	5	07:30	0	0
12/08/2010	4	00:47	0	00:47	4	0
13/08/2010	4	04:01	2	01:54	2	0
14/08/2010	2	05:27	2	01:48	0	0
15/08/2010	2	00:31	2	04:20	0	0
16/08/2010	3	03:21	3	02:20	0	0
17/08/2010	2	05:39	2	05:50	0	0
18/08/2010	1	01:25	2	03:07	0	1
19/08/2010	1	01:33	2	03:13	0	1
20/08/2010	0	03:15	1	07:55	0	1
21/08/2010	0	03:59	0	03:59	0	0
22/08/2010	0	00:07	0	00:07	0	0
23/08/2010	0	02:21	0	02:21	0	0
24/08/2010	0	05:41	1	08:54	0	1
25/08/2010	1	00:51	0	00:51	1	0
26/08/2010	2	06:39	2	06:21	0	0
27/08/2010	1	01:55	1	06:30	0	0
28/08/2010	1	05:47	1	04:10	0	0
29/08/2010	2	01:11	2	03:48	0	0
30/08/2010	2	02:51	2	02:28	0	0
04/09/2010	3	12:13	3	14:13	0	0
05/09/2010	4	06:19	4	07:15	0	0
06/09/2010	2	02:07	1	04:33	1	0
07/09/2010	0	21:20	1	14:45	0	1
11/09/2010	0	00:12	0	00:12	0	0
12/09/2010	1	08:05	2	09:12	0	1
13/09/2010	3	00:13	1	03:50	2	0
20/09/2010	3	07:59	2	07:23	1	0
22/09/2010	2	04:49	2	05:16	0	0
23/09/2010	2	14:11	2	09:10	0	0
24/09/2010	2	03:15	2	03:47	0	0
25/09/2010	2	03:25	2	03:50	0	0
26/09/2010	2	03:25	3	04:13	0	1
27/09/2010	3	03:27	3	03:40	0	0
29/09/2010	3	04:21	3	03:12	0	0
30/09/2010	4	00:47	4	01:35	0	0
Total	89		84		12	7

[18] FRR: the number of sunspots detected by the automatic method but not by the NOAA

FAR: the number of sunspots detected by the NOAA but not by the automatic method.

The first column in Table 7, from left to right, shows the date SODISM images were captured, followed by the number of sunspots detected by the automatic method. The third column shows the time of capture for each image, the fourth column shows the number of sunspot in the NOAA catalogue, the fifth column is the time indicated in NOAA catalogue and the final two columns are FRR, representing the number of sunspots detected by the automatic method but not by the NOAA, and FAR, representing the number of sunspots detected by the NOAA but not by the automatic method

The total number of sunspots detected is listed in the last row in Table 7 for the automated and NOAA methods. The recognition rate is calculated as follows:

$$Recognition\ Rate = \frac{sum\ of\ sunspots\ detected\ by\ the\ automatic\ method\ -\ sum\ of\ the\ FAR}{sum\ detected\ by\ NOAA}$$

$$\frac{89\ -\ 7}{84} \times 100 = 97.6\%$$

3.9 Visual sunspot verification

In order to verify the detection results, a few randomly selected images of recognized sunspots were chosen. The detected areas are magnified and compared with the original images and Figure 75 shows the results.

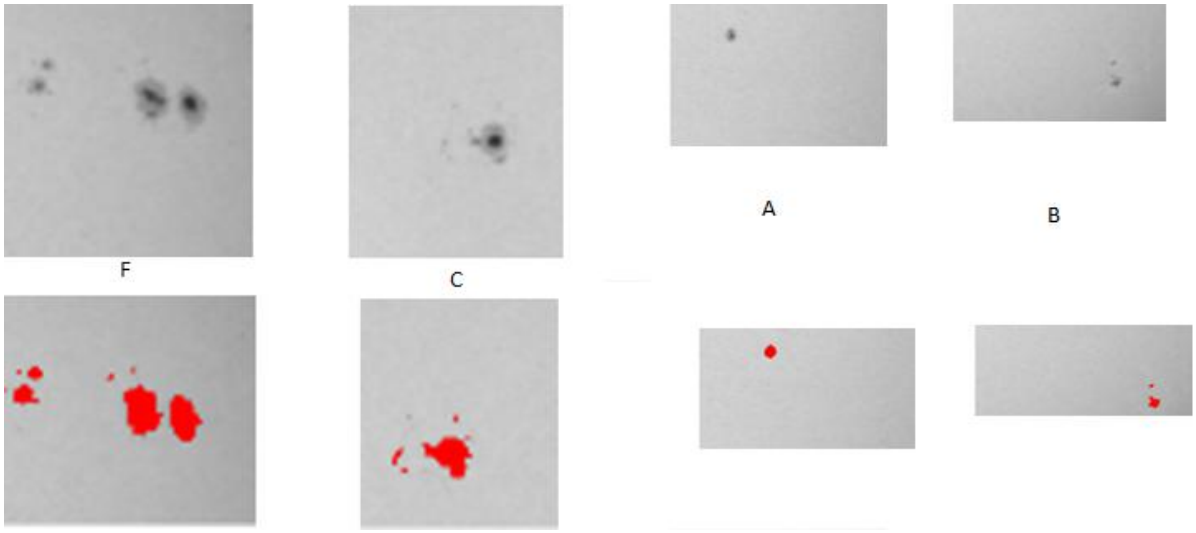


Figure 75: Enlarged most recognised sunspots: the top images show the original sunspot images and the bottom images show the zoomed images, according to MacIntosh class (A) unipolar no penumbra, class (B) bi-polar no penumbra, class (C) bi-polar with penumbra and class (F) bi-polar penumbra on spot

The sunspot detection and grouping algorithm was tested on the SODISM archived images for the period August 2010-January 2014 for 535 nm. Figure 76 shows the comparison with the NOAA catalogue, and the 95.17%. Correlation accuracy achieved. Figure 77 shows the plot.

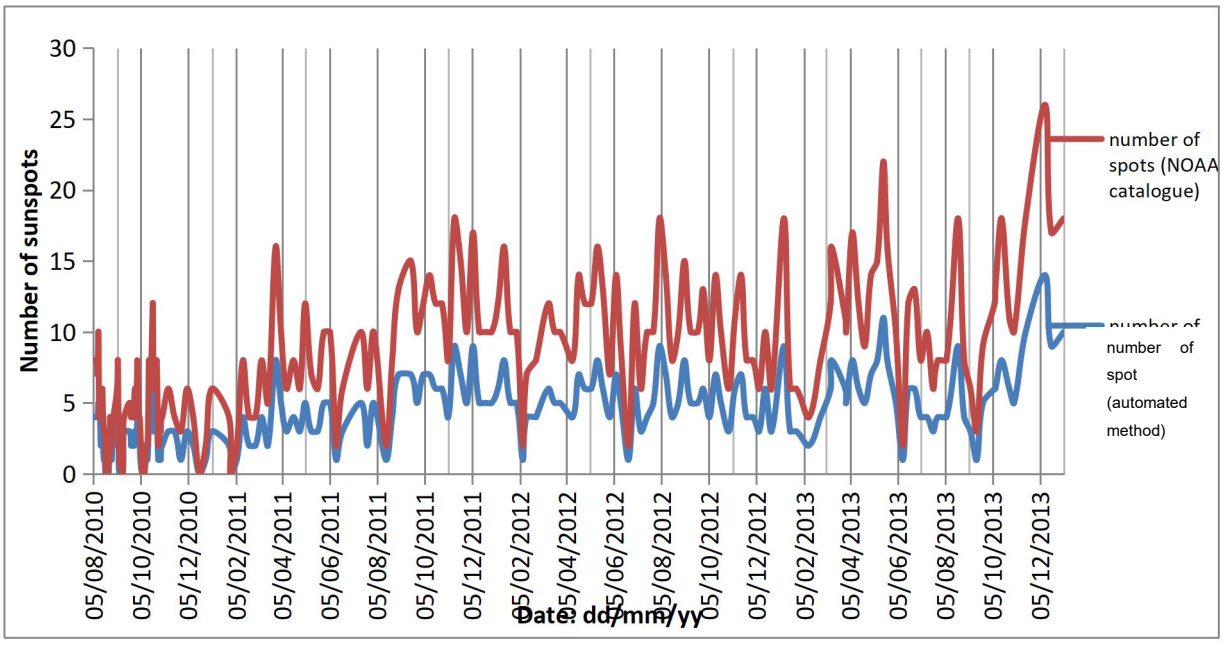


Figure 76: USAF/NOAA sunspot numbers vs SODISM sunspot numbers by automated method

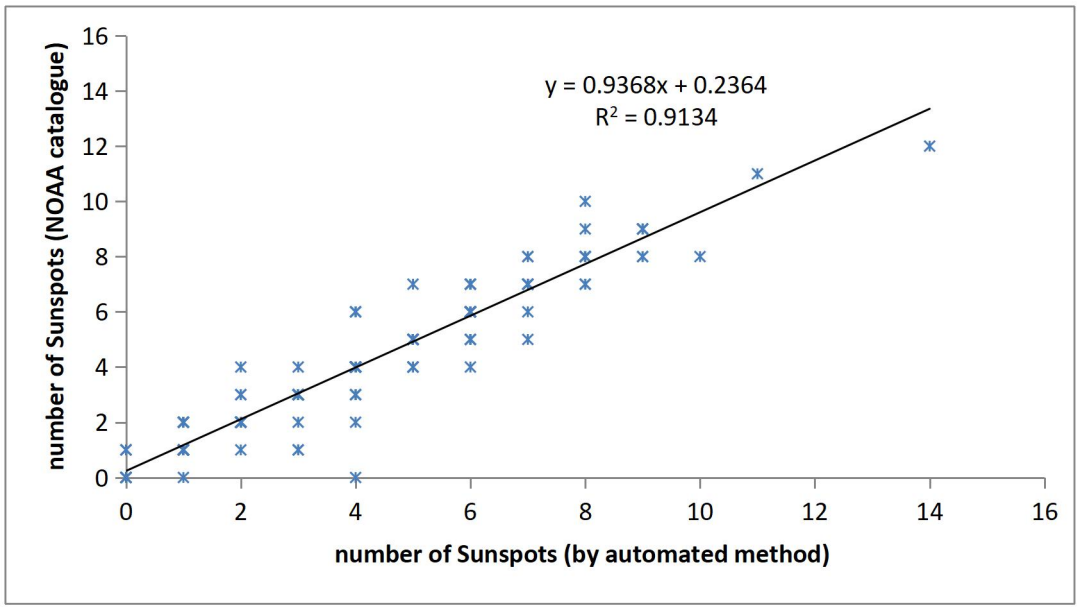


Figure 77: Showing the plot of sunspots vs time for wavelength 535 nm between 2010-2014

The sunspot areas calculated by the automated sunspot detection method were compared to those recorded by the NOAA. More than two hundred individual images were compared from August 2010 to January 2014, and results are illustrated in Figure 77. Excepting the regions with images from 2013 and 2014, where there was degradation in the images, there is a strong correlation between the automated sunspot area and the USAF/NOAA sunspot area. A visual inspection of the SODISM images seems to confirm that the automated procedure works well and confirms the results.

3.10 Computation sunspot area

The area of the whole sunspot, including Umbra and Penumbra, has been calculated in this research using images from SODISM after determining and verifying the sunspots. Generally, the area is the product of a number of pixels, and the area of each pixel the area of a pixel is the square of the pixel size (Pucha, Hiremath and Gurumath, 2016). Equation 13 calculates pixel size:

$$\text{Pixel size} = \frac{\text{Radius of sun in arcseconds}}{\text{Radius of sun in pixels}} \quad (13)$$

Calculating the number of pixels in the whole sunspot is then estimated and multiplied by the area of the pixel to obtain the sunspot area. This area will

be expressed in a millionth of an area of a solar hemisphere, which is the same unit in NOAA catalogue, and it is the standard unit used to measure the area of sunspots. The proposal FF catalogue will contain the area of the sunspot for each day at a specific time in million square kilometres.

3.11 Bright region segmentation

Bright regions in SODISM images include active regions, faculae and networks (Meftah et al., 2016), The difficulties in SODISM images arise because these areas cannot be separated on account of their low resolution Moreover, the noise will affect the ability to find a suitable threshold for bright region detection. Nevertheless, many experiments have been applied on SODISM images in different wavelengths with little or no success, with the exception of wavelength 393 nm, which gives limited success with the suggested algorithm, depending on an estimated automatic threshold if tested and corrected to obtain maximum possible accuracy. The FITS images header files contain information on the position and size of the solar disk. Unfortunately ., this information is not precise; investigation shows the discrepancies amount to several pixels, varying in magnitude from image to image, and before applying the proposed method, the algorithm 3 to determine solar disk edges. Once that happens the following algorithm 7 is applied to obtain bright region segmentation.

Algorithm 7: Bright region segmentation

i. Convert the image scale from signed 32 bit to unsigned 8 bit; this helps to make them easy to handle.

ii. Remove unwanted noise by applying the Kuwahara Filter (figure 78(a)).

iii. Apply limb darkening removal to distribute the intensity overall the solar disk (figure 78(b)).

iv. Outliers are removed with a threshold of 4 to remove white regions (figure 78(c)). Subtract filtered image in figure 78(c) with the de-noised image from figure 78(a). (the result is illustrated in Figure 78(d)) then apply erode operation.

v. Apply filter “Multi Thresholder Max Entropy” to operate on the whole stack histogram to determine the threshold and detect bright regions as shown in Figure 78(e).

vi. Make a visual comparison, apply a binary overlay in red, and superimpose the original image. Figure 78(f) shows sample result.

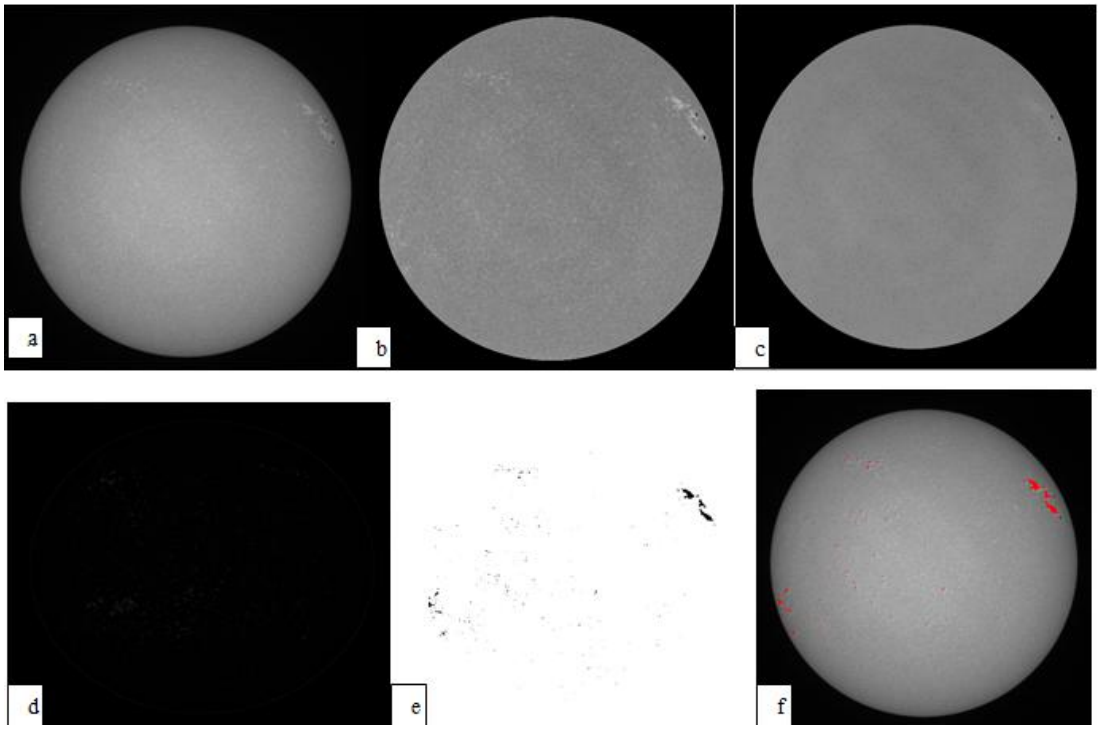


Figure 78: Figure 78: Image at wavelength 393 nm at 26th January 2011 at 11:11 (a) Original image disturbed by instrument noise; (b) Clean image after Kuwahara filter; (c) Filtered image without bright regions; (d) Gradient on the image Binary image showing bright candidates; (e) Recognised bright regions after MultiThresholder MaxEntropy; (f) Overlay original image with bright regions detected

The algorithm 7 has been applied to approximately 300 images on 393 nm and they have been visually inspected. The images were captured in the last months of 2012, and between 2013-2014 and the threshold must be manipulated manually to detect bright regions. That may be caused by raised noises at the end life of the PICARD Satellite and degradation caused by deterioration in lens quality at the end of its useful life.

3.11.1 Bright region verification

For each bright region on the surface of the sun detected, a small threshold of 0.1 is applied to set the growth of the region and this growth is compared with all unallocated neighbouring pixels. The method in algorithm 7 uses the difference between a pixels intensity value and the region's mean to determine similarity. Therefore pixels measured in this manner with the smallest difference will belong to the region but when the intensity difference between the region average and the new pixel grows larger than the entered threshold or if there is a notable increase in the volume of detected pixels the

process will cease. The threshold will be increased by the smallest value 0.01 repeating the process to detect pixels of the biggest region possible.

The first results were assessed visually and showed matching between bright regions detected and bright regions in the original images in 2010, 2011 and early months of 2012. The method then increased the threshold manually to get acceptable results. The second assessment was made by calculating the FF for bright regions. Unfortunately, the lack of reference sources to compare the proposed method results in the same time period meant the results were unreliable unless further research can calculate FF in the same time period and compare it with our results.

3.12 Conclusion

In this chapter a method to enhance the image quality and resolution has been proposed, and an automatic image segmentation pipeline has been developed. Accordingly, the combination of these two algorithms, with their morphological and image differentiating operations, provides an automatic sunspot detection system. A thorough evaluation of each component within the detection system has been conducted on hundreds of images. To further explore the method, two different image formats (JPG and FITS) of the SODISM data acquired from PICARD were used and images in both of these forms have been successfully segmented to detect sunspots. Enhanced methods were also applied to the JPG images before the segmentation methods were applied.

Image enhancement was first explored to improve the visibility of solar features, such as sunspots, in images at five wavelengths. A combination of Discrete and Stationary wavelet decomposition was used for enhanced resolutions of the SODISM images. Many images were used to compare the performance of the proposed technique at different wavelengths with that of various established interpolation techniques, both for edge detection and edge integration (Kimmel and Bruckstein 2003). It is clear from Figures 30-39 that the images enhanced by the proposed technique are sharper and have been visually improved compared to the results of the other tested methods. Quantitative comparisons, using the calculation of SSIM, PSNR and MSE values, also confirm the effectiveness of the proposed method for the

enhancement of resolutions. Besides that, the comparison of NAE, NCC and SC results supports the viewpoint that the proposed method is powerful, and gives better results than the bicubic method. In all cases, the proposed technique has been subjected to tests to evaluate image enhancement quality (SSIM) to compare the original images with compressed images (MSE) and to check the ratio of the maximum power of a signal to the power of distorting noise that affects the quality of representation (PSNR) to establish and improve visual clarity.

As a key technique for identifying regions of interest, image segmentation was similarly explored through an evaluation and comparison of the results from the NOAA catalogue. In general, the system described in this chapter has been shown to be promising; of 84 sunspots, 89 were automatically detected, which was a little more than the NOAA method. In some cases, this might be due to the different times. Thus, the proposed system matches the detected $\approx 98\%$ of the sunspots for the period August 2010- Sep 2010; similarly, a comparison for period August 2010 to January 2014 shows a correlation coefficient of 95.17%. Moreover, the sizes of sunspots are matching 93% in NOAA catalogue (figure 97); the appeal of the segmentation proposal method is that it is appropriate to recognize sunspots for lower resolution SODISM images, particularly the images associated with instrument noises. Despite the degradation in the quality of SODISM images through their lifetime, the automatic detection of sunspots in SODISM images has been shown to be possible with suitable pre-processing procedures. Meanwhile, the method proposed by Meftah et al. in 2016, applied to 393 nm SODISM images required manual interaction to optimize the threshold. This is calculated using the Otsu method, whereas the proposed method in this research is based on a histogram of all the interior solar disc pixels. Furthermore, the approach in this study does not involve manual interaction and iterative optimization steps.

Finally, the automated methods applied to detect bright regions were successful at 393 nm in 2010, 2011 and the early months of 2012.Unfortunately, the method does not work for images captured in the final months of 2012 and for 2013 and 2014, due to the increasing volume of

noise in SODISM images and the mixing of bright regions with unwanted noises.

It is a serious weakness that there is no easily accessible comparative reference system to assess the bright regions of SODISM images other than visual assessment. This is because the bright regions in SODISM images which include faculae, networks and active regions are difficult to segment [ ] while images from telescopes such as the SOHO satellite areas are more easily separated. SOHO has two kinds of images (intensity-grams and magnetograms) making it easier to distinguish each area. The results of bright regions obtained from the proposed method in algorithm 7, appear visually promising, and can be used for comparative purposes in future research.

CHAPTER FOUR

4. Filling factor

4.1 Introduction

The calculated filling factor (FF) for a feature reflects the fraction of the solar disc covered by the feature and the assignment of reference synthetic spectra. Moreover, it can be applied to the modelling of the Solar Spectral Irradiance (SSI). When discussing imaging sensor arrays, the filling factor is defined as the ratio of a pixel's light sensitive area to its total area (Rousseeuw and Hubert, 2011). Similarly, when studying solar sunspots, the value of the FF represents the ratio of a magnetized component in each observed pixel (Rousseeuw and Hubert, 2011). Research has shown that the umbrae and penumbrae generally have large FFs close to unity, and it can be inferred that the magnetized component almost entirely covers the pixels (Rousseeuw and Hubert, 2011) In comparison, granules outside sunspots show small FFs because of the dominance of non-magnetized gas.

For this study, the FF is calculated as a function of the radial position on the Sun's disc. The solar disc is divided into 11 concentric rings (Ashamari et al., 2015) starting with an inner radius (RI) and an outer radius (RO). In Tables 8 and 9, these radii are defined as relative to the radius of the solar disc. In the case of SODISM images, the last interval would be between 0.95 and the boundary of the solar disc. If a sunspot is dissected by a concentric circle, the number of pixels within the sunspot crossed by one of those concentric rings will be calculated based on the area dissected between the two adjacent regions. This chapter focuses on the computation of FF. As stated above, the value of the FF gives an indication as to whether the region under consideration has large or small sunspots. Furthermore, these values can be used to compare sensors. For instance, this study uses images from the Solar and Heliospheric Observatory (SOHO) for comparative purposes. Thus, the chapter presents in detail the computation of FF. Firstly, the solar radius and centre of the solar disc are acquired where, for the FITS images, this forms part of the meta-information provided in the header. In comparison,

for the JPG images, image-processing techniques are employed to obtain the necessary information. Towards the end of this chapter, a novel cataloguing technique is proposed and implemented. Previous researchers have failed to give a clear picture of the whole dataset by only reporting the FF for a few sample images. Thus, parametric and non-parametric statistical measures are used to evaluate the experiments conducted on the whole image dataset.

4.2 FF computation

As previously mentioned, the inner and outer radii are calculated, as shown in Tables 8 and 9. FFs have been calculated for the 535 nm wavelength in the FITS format, and in four other wavelengths in the JPG format. The results have been divided into two main parts; the first part is for SODISM 535 nm, 607 nm and 782 nm images in FITS format and these results are compared with the FF of SOHO images for the same period. The second section is for the four wavelengths in a JPG format in which the results for all enhanced images are compared with the benchmark images. Table 8 presents a sample of the radii for both images format, and figure 79 shows an example of basic idea of FF solar disc divided by 11 concentric rings.

Table 8: Relative radius values

Index	Inner radius (relative radius)	Outer radius (relative radius)
1	0.00	0.07
2	0.07	0.16
3	0.16	0.25
4	0.25	0.35
5	0.35	0.45
6	0.45	0.55
7	0.55	0.65
8	0.65	0.75
9	0.75	0.85
10	0.85	0.95
11	0.95	1.05

Table 9: Example of the numbers of sunspot in the different parts of the solar disc in images taken on 5th August 2010, 04:49.

Index	Inner radius	Outer radius	FF
1	0	0.07	0
2	0.07	0.16	0
3	0.16	0.25	0
4	0.25	0.35	0
5	0.35	0.45	11.70×10^2
6	0.45	0.55	0
7	0.55	0.65	0
8	0.65	0.75	0.1855×10^2
9	0.75	0.85	0.9445×10^2
10	0.85	0.95	3.38×10^2
11	0.95	1.05	0
Total of Pixels			1621

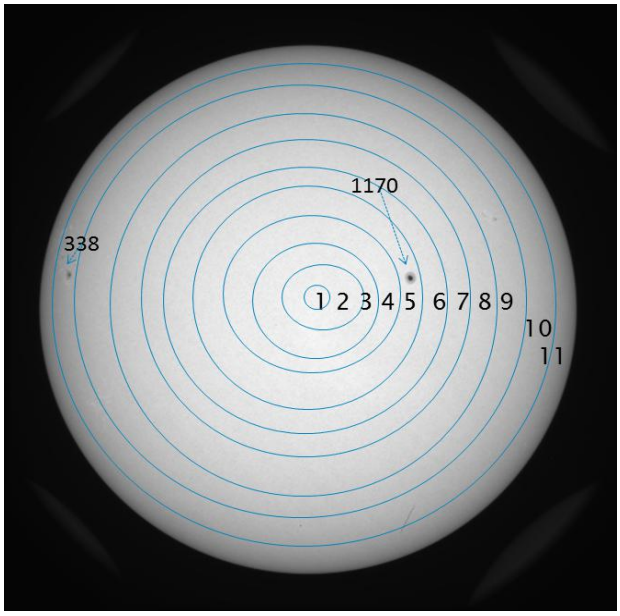


Figure 79: SODISM image (wavelength 535 nm) taken on 5th August 2010, 04:49 with FF value

4.2.1 Filling factor for SODISM images in FITS format

The previously outlined segmentation procedure has been applied to the entire set of downloaded 535 nm image data in order to detect the sunspots and calculate their FFs. Figure 85 highlights the procedure for the calculation of the FF for data collected from 5th August 2010 to 25th December 2010, and Figure 81 shows the results of sunspot FF calculations for data

collected from August 2010 until the end of the life of the PICARD satellite on 1st January 2014. Furthermore, data collected from August 6th 2010 to December 19th 2010 are compared to the data collected from SOHO also retrieved from the same period, and Figure 80 shows the results. The correlation coefficient calculated for the two data sets was found to be 0.985. In order to perform an extensive evaluation, the FFs were further computed for all the images retrieved over the four-year period of PICARD's lifetime, and the results are shown in Figure 82.

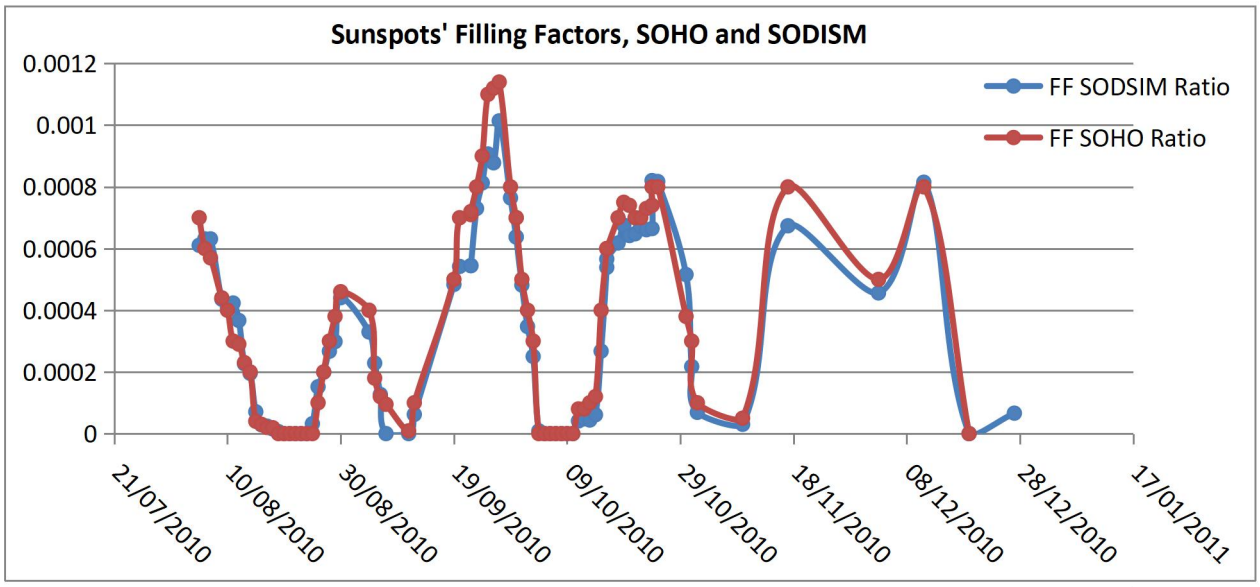


Figure 80: Comparison of sunspot FFs calculated from SOHO and SODISM 535 nm images from 5th August 2010 until 25th December 2010

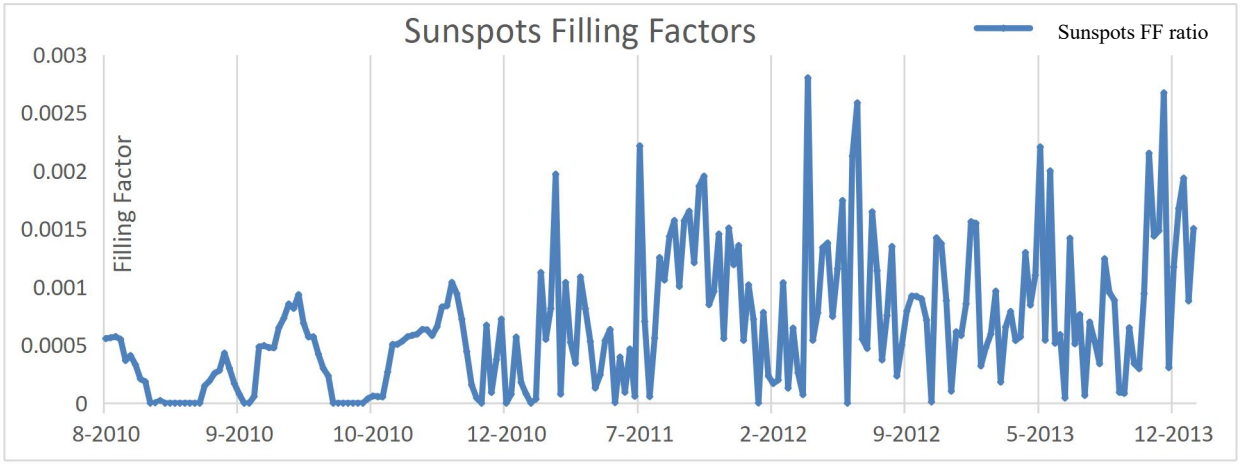


Figure 81: Sunspot FFs calculated from SODISM 535 nm images over the lifetime of the satellite.

The distribution of the sunspot FFs over the period from August 2010 to January 2014 is shown in Figure 82.

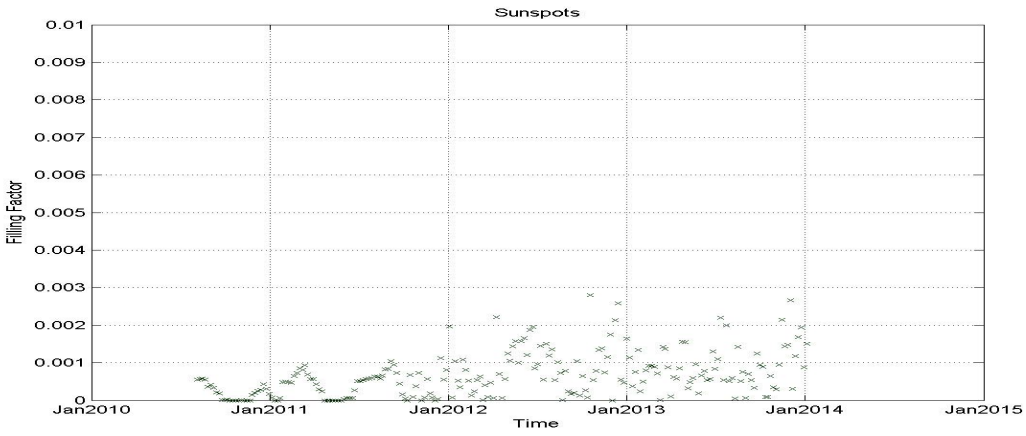


Figure 82: Area coverage of FFs for sunspots

The last data set was obtained from the French National Centre for Scientific Research (FNCSR) for a wavelength of 607 nm, from 22 September 2010 to 1st January 2014; the FF for this data was calculated and compared with that from SOHO for the period September to December 2010. The matching correlation was 98.9%, which demonstrates that the method introduced here gives excellent results for the 607 nm wavelength. Figure 83 shows the comparison with the SOHO satellite results.

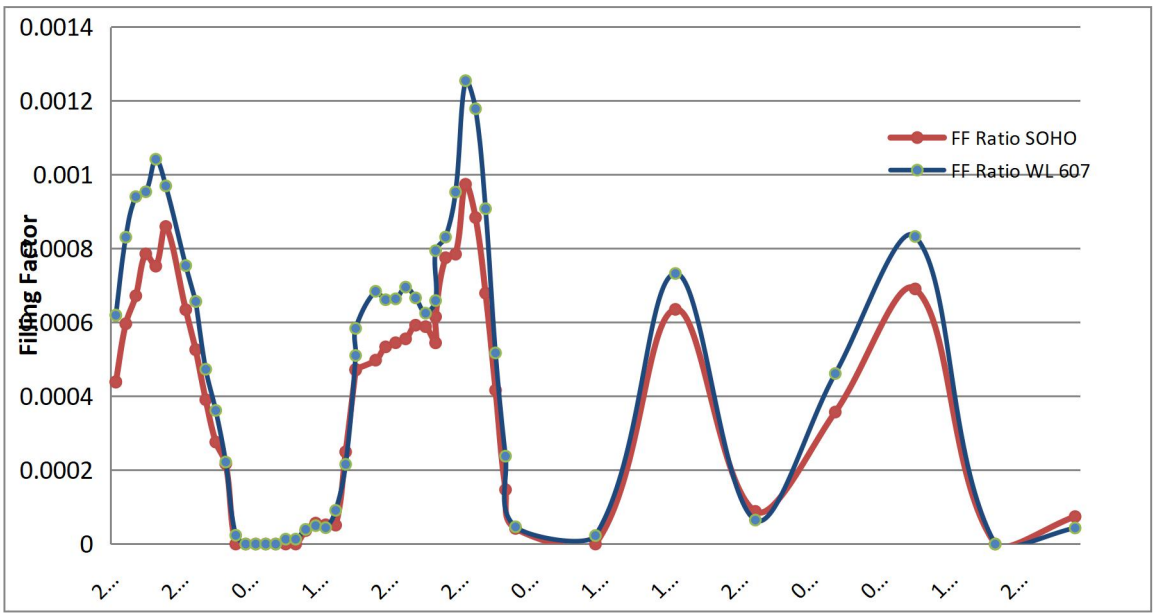


Figure 83: Comparison of the sunspot FF ratio calculated from SOHO and SODISM images from 22nd September 2010 until 25th December 2010

For the SODISM 782 nm images, a correlation of 93% with SOHO was achieved for the same period (from 22nd September 2010 to 19th December 2010), and Figure 84 shows this comparison.

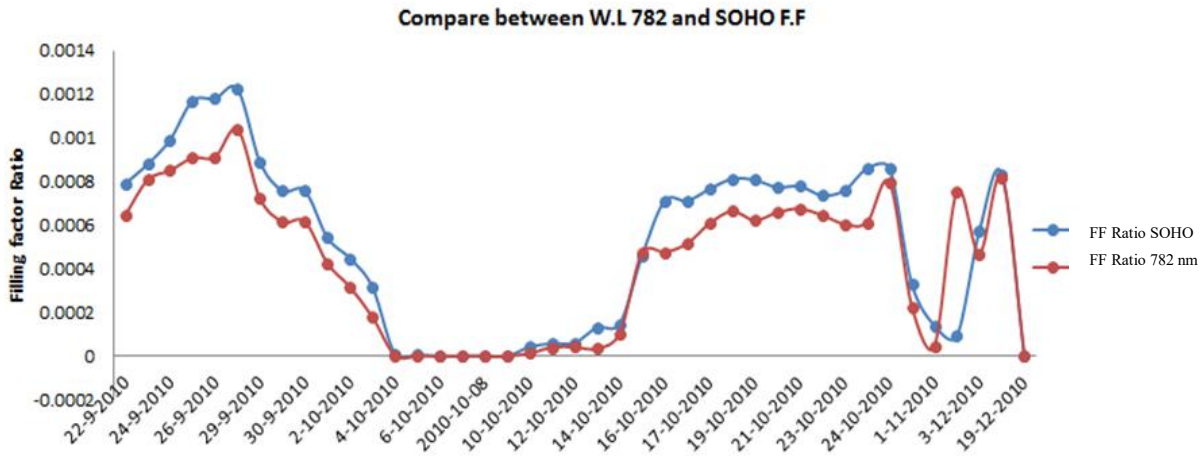


Figure 84: A comparison of sunspot FFs ratio calculated from SOHO and SODISM images from 22nd September 2010 until 25th December 2010

4.2.2 Filling factor JPG format

This section presents the results of the sunspot segmentation, and the calculated percentage improvement in resolution, after applying the image quality enhancement method. Due to the image capture conditions, images may appear blurred and become candidates for improvement before visual assessment; this was the case with some of the images downloaded from SODISM. As such, the FFs are calculated for five wavelengths before and after the application of the quality enhancement technique. Again, the steps taken to calculate the FF are outlined diagrammatically in Figure 85.

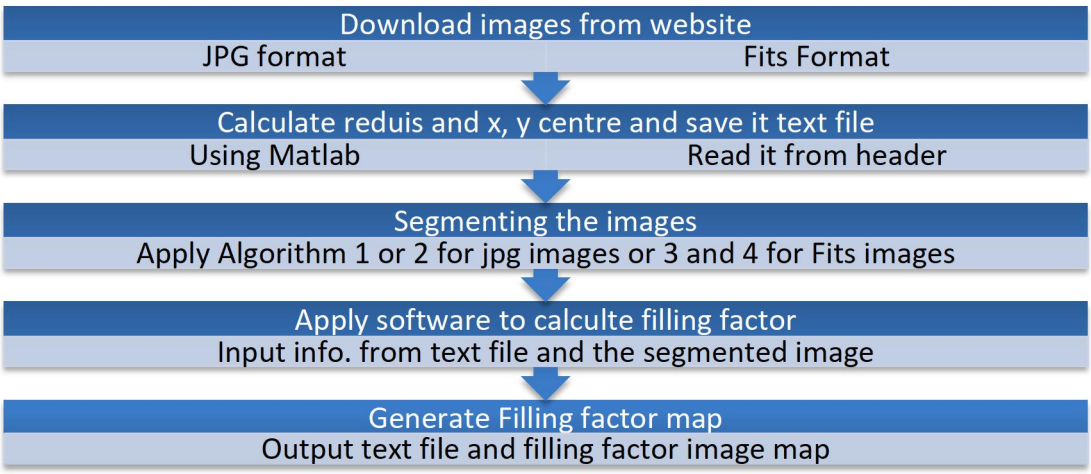


Figure 85: Block diagram for general steps used to calculate the FF

The generation of the FF map requires two input files; the first is a text file that contains the radius and the X, Y coordinates of the centre of the solar disc, and the second file is the segmented image. Figure 86 shows the basic ideas and calculation procedure of the radius and the X, Y centre.

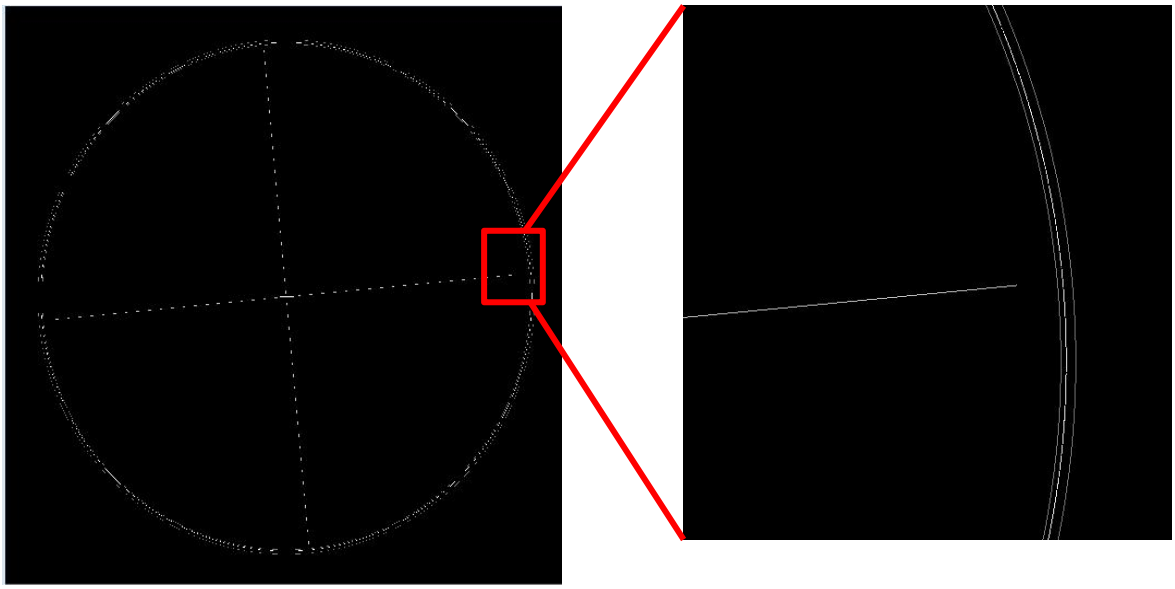


Figure 86: Radius and X, Y centre calculation method

As shown in Figure 86, since the shape of the solar disc is not completely circular, the average of the boundaries (middle line) is computed by generating the outer and inner boundaries of the solar disc to obtain an approximation of the circular shape. Thereafter, the radius of the shape is calculated as the distance between the centre of the solar disc and the middle line created above.

4.2.2.1 Wavelength 215 nm:

As previously mentioned, image quality enhancement is first conducted to eliminate noise and unwanted artefacts, and Figure 87 shows a sample input image at a wavelength of 215 nm. Although the image quality is poor at this wavelength, the results of the SSIM, MSE and PSNR computed before and after applying the proposed enhancement algorithm show clear improvements. Table 10 shows an example for improving the value of PSNR and SSIM wavelength 215 nm at 3 different dates compared to the bicubic method. .

Table 10: Comparison of PSNR and SSIM values for 215 nm

W.L	Date and Time	PSNR		SSIM		% PNSR	%SSIM
		Bicubic	DWT	Bicubic	DWT	DWT:BIC	
215	2011 Oct.11 at 04:05	3452×10^{-2}	4058×10^{-2}	67.1×10^{-2}	80.6×10^{-2}	14.91%	16.7%
	2010 Oct. 01 at 03:21	3786×10^{-2}	4390×10^{-2}	69.1×10^{-2}	78.5×10^{-2}	13.76%	11.9%
	2014 Jan. 01 at 1644	3764×10^{-2}	4368×10^{-2}	65.1×10^{-2}	73.6×10^{-2}	13.83%	11.5%

The application of Peak Signal to Noise Ratio shows an improvement in image clarity of approximately 14% while the average of Structural Similarity Index Metric has increased by nearly 13% supporting our previous conclusion that it represents an effective enhancement method, although this improvement in the resolution does not extend to proposed segmentation methods with a wavelength of 215 nm.

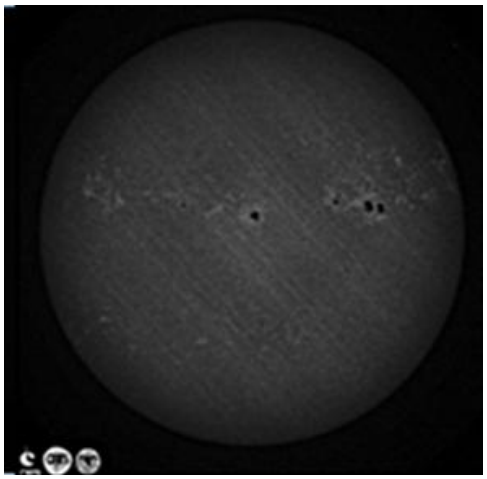


Figure 87: SODISM image (wavelength 215 nm) taken 1st October 2011 at 02:55

4.2.2.2 Wavelength 393 nm:

For this wavelength, the proposed enhancement algorithm was first used to improve the image quality. Thereafter, segmentation was performed, as shown in Figures 88, 89 and 90. A numerical comparison of the segmentation results is shown in Table 10. The solar disc is divided into concentric rings in a manner similar to previous studies (Ashamari et al. 2015; Pradels et al. 2008). The ratios of the areas of sunspots within each ring to the areas of the corresponding ring, called the sunspot-FF, are shown in a separate column for every image. Also shown in Table 9 are the inner and outer radii of each ring in pixels, RI and RO; these are expressed as a ratio

of the solar radius, the total number of segmented sunspot pixels, and the solar radius for each image.

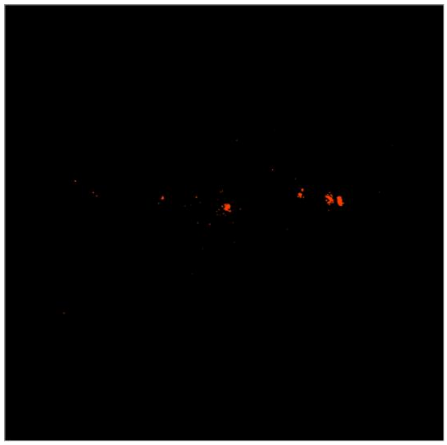


Figure 88: Sunspots detected in benchmark image at wavelength 393 nm taken on 1st October 2011 at 02:47

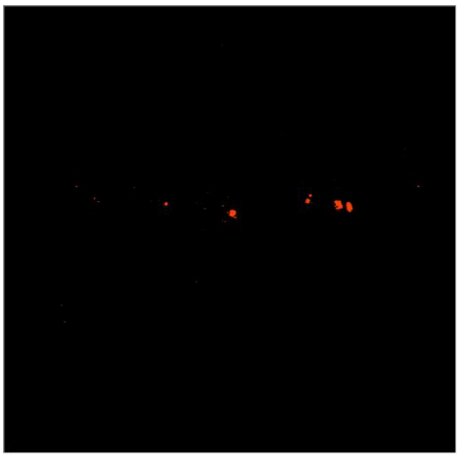


Figure 89: Sunspots detected in enhanced image at wavelength 393 nm taken on 1st October 2011 at 02:47

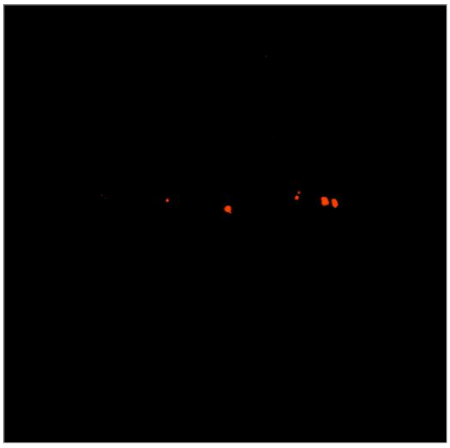


Figure 90: Sunspots detected in test image at wavelength 93 nm taken on 1st October 2011 at 02:47

Table 11 shows the results of images taken at 1st October 2011, where RI and RO represent the internal and external radius for each ring, and their FF ratio for test, enhanced and benchmark images, the table also shows the radius of test, enhanced and benchmark images and the over all of FF shows in last row.

Table 11: Sunspot FFs for the solar image taken on 1st October 2011 at wavelength 393 nm are shown as percentages for the test, enhanced and benchmark images in the labelled columns.

Index	RI	RO	Test	Enhanced	Benchmark
1	0.00	0.07	410×10^{-2}	359×10^{-2}	437×10^{-2}
2	0.07	0.16	66.3×10^{-2}	102×10^{-2}	107×10^{-2}
3	0.16	0.25	0	3.2×10^{-2}	6.8×10^{-2}
4	0.25	0.35	11.8×10^{-2}	12.5×10^{-2}	11×10^{-2}
5	0.35	0.45	18.9×10^{-2}	27×10^{-2}	28×10^{-2}
6	0.45	0.55	47.2×10^{-2}	5.7×10^{-2}	8.4×10^{-2}
7	0.55	0.65	31.1×10^{-2}	65.9×10^{-2}	60×10^{-2}
8	0.65	0.75	0.5×10^{-2}	1.9×10^{-2}	1.5×10^{-2}
9	0.75	0.85	0	0.9×10^{-2}	1.3×10^{-2}
10	0.85	0.95	0.3×10^{-2}	0.6×10^{-2}	0
11	0.95	1.00	0	3.2×10^{-2}	1.2×10^{-2}
Number of feature pixels			0.9×10^{3}	4.132×10^{3}	$4{,}107\times10^{3}$
Solar radius in pixels			0.45×10^{3}	0.898×10^{3}	0.898×10^{3}
Overall FF			141.5×10^{-3}	162.2×10^{-3}	161.2×10^{-3}

From an inspection of the results in Table 11, the most obvious difference is that the overall FF percentage for the enhanced image is closer to the benchmark image than the test image. The size of test images is 512x512 pixels, so a comparison with enhancement and benchmark images should multiply the test image value by a factor of four. In the individual rows, the results for the enhanced images are closer to the benchmark than the test image in 7 out of 11 cases, there is no significant difference in one, and are closer to the test image than the benchmark in three cases. The biggest discrepancy is in the first row, but in reviewing the entire table, the enhanced results are superior.

4.2.2.3 Wavelength 535 nm:

Following the same procedure applied to the 215 nm and 393 nm SODISM images, the images at the 535 nm wavelength were segmented, as illustrated in Figures 91, 92 and 93 for benchmark, enhanced and test images respectively. Furthermore, Table 12 presents numerical values of the FF computation.

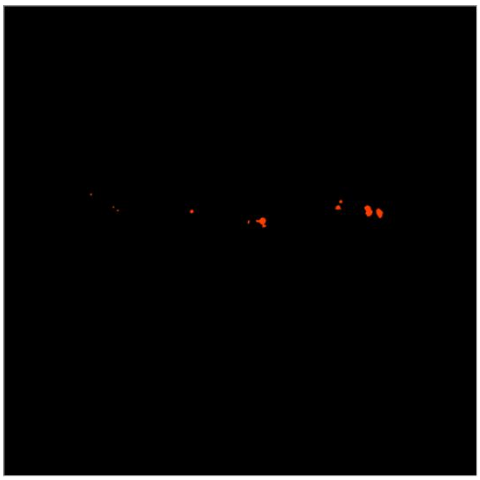


Figure 91: Sunspots detected in benchmark image at wavelength 535 nm taken on 1st October 2011 at 12:30

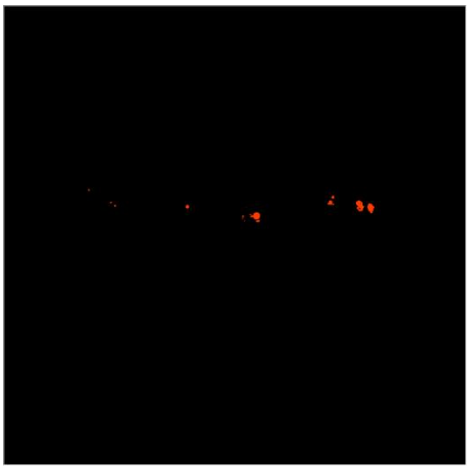


Figure 92: Sunspots detected in enhanced image at wavelength 535 nm taken on 1st October 2011 at 12:30

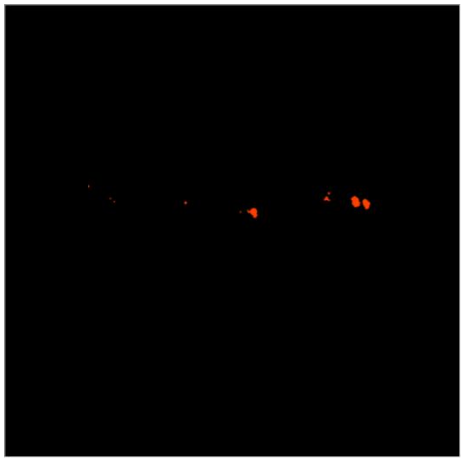


Figure 93: Sunspots detected in test image at wavelength 535 nm taken on 1st October 2011 at 12:30

Table 12 is similar to table 11 in that it shows the same information with different wavelength and the same image date in that it shows the results of images taken on 1st October 2011. The table also shows the radius of test, enhanced and benchmark images, and the overall FF of the three images in the bottom row. An inspection of the results in Table 12 indicates clearly that the overall FF percentage for the enhanced image is closer to the benchmark image than it is to the test image. This is reflected in the number of features if the test value is multiplied by a factor of four to account for the different radii. In the individual rows, the results for the enhanced image are closer to the benchmark than the test image in five of 11 cases, are the same (at zero) in five cases and are marginally closer to the Test in one.

Table 12: Sunspot FFs for the solar images taken on 1st October 2011 at wavelength 535 nm are shown as percentages for the test, enhanced and benchmark images.

Index	RI	RO	Test	Enhanced	Benchmark
1	0.00	0.07	0	0	0
2	0.07	0.16	235.6×10^{-2}	222.7×10^{-2}	210.5×10^{-2}
3	0.16	0.25	0	0	0
4	0.25	0.35	7.3×10^{-2}	12.9×10^{-2}	12.4×10^{-2}
5	0.35	0.45	0	0	0
6	0.45	0.55	16.2×10^{-2}	19×10^{-2}	19.7×10^{-2}
7	0.55	0.65	43×10^{-2}	39×10^{-2}	39.8×10^{-2}
8	0.65	0.75	37×10^{-2}	26.5×10^{-2}	25×10^{-2}
9	0.75	0.85	1.2×10^{-2}	1.2×10^{-2}	1.5×10^{-2}
10	0.85	0.95	0	0	0
11	0.95	1.00	0	0	0
Number of feature pixels			1.114×10^{3}	4.052×10^{3}	3.975×10^{3}
Solar radius in pixels			0.4498×10^{3}	0.9008×10^{3}	0.9008×10^{3}
FF %			17.5×10^{2}	15.9×10^{2}	15.6×10^{2}

The number of features in table 12 shows that the enhanced image is 4052 pixel and the benchmark is 3975 pixel which is almost the same value, while the test image is 1114 pixel, it is clear that the enhanced image still keeps the same features and details in the original benchmark image . Moreover, the filling factor in the enhanced image (159.1×10^{-3}) is closer to the benchmark image (156.1×10^{-3}) than the test image (175.2×10^{-3}) which again illustrates the superior enhanced images.

4.2.2.4 Wavelength 607 nm:

There are 126 images in JPG format, each with a size of 2048×2048 pixels. Figure 94 shows a sample of the input image, and Figure 94 shows the FF of the corresponding image.

Figure 94: (left) Input image for 607 nm at 4th October 2010 at 06:41, (right) Sunspot detected

Table 13 shows the FF for wavelength 607 nm images taken on 4th October 2010 at 6:41, the last column shows the sunspots in each ring and the total of sunspot area/pixel shows in the last row. This example has been applied to all available images to calculate the FF.

Table 13: Example of numbers of sunspot features in the different parts of the solar disc in images taken on 4th October 2010 at 06:41

Index	Inner radius	Outer radius	Strip area/pixel	Sunspot area/pixel
1	0	0.07	1.2×10^{4}	0
2	0.07	0.16	5.2×10^{4}	0
3	0.16	0.25	9.3×10^{4}	0
4	0.25	0.35	1.5×10^{5}	400×10^{-2}
5	0.35	0.45	2.0×10^{5}	100×10^{-2}
6	0.45	0.55	2.5×10^{5}	0
7	0.55	0.65	3.0×10^{5}	100×10^{-2}
8	0.65	0.75	3.5×10^{5}	0
9	0.75	0.85	4.0×10^{5}	0
10	0.85	0.95	4.5×10^{5}	0
11	0.95	1	2.4×10^{5}	10600×10^{-2}
Total of sunspots				112

4.2.2.5 Wavelength 782 nm:

Figure 95 shows a sample of the input image and the segmentation output with the FF for an example of an image at wavelength 782 nm.

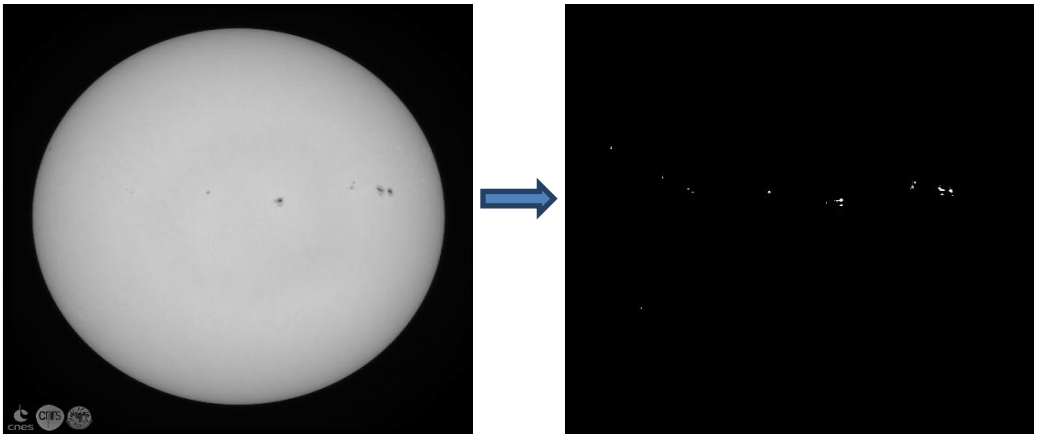


Figure 95: Input image at wavelength 782 nm taken on 1st November 2011 at 22:09

Table 14 illustrates the FFs of the wavelength 782 nm images taken on 1st November 2011 at 22:09; this method has been applied to all available images to calculate their FFs

.

Table 14: Example of numbers of sunspots in the different parts of the solar disc in images taken 1st November 2011 at 22:09

Index	RI	RO	Strip area/pixel	Sunspot area/pixel
1	0	0.07	1.2×10^4	0
2	0.07	0.16	5.2×10^4	0
3	0.16	0.25	9.3×10^4	5.2×10^2
4	0.25	0.35	1.5×10^5	0
5	0.35	0.45	2.0×10^5	0
6	0.45	0.55	2.5×10^5	0.32×10^2
7	0.55	0.65	3.0×10^5	2.6×10^2
8	0.65	0.75	3.5×10^5	5.9×10^2
9	0.75	0.85	4.0×10^5	0.78×10^2
10	0.85	0.95	4.5×10^5	0
11	0.95	1	2.4×10^5	0.33×10^2
Total of sunspots				1520

4.2.3 Filling factor bright regions

The FF for bright regions as in the FF sunspot procedure will require the same factors. This means that the xy centre and radius will be calculated for each image and as previously mentioned, the solar disk will be divided into 11 rings. Each area will be calculated and figure 96 shows an example for bright regions FF.

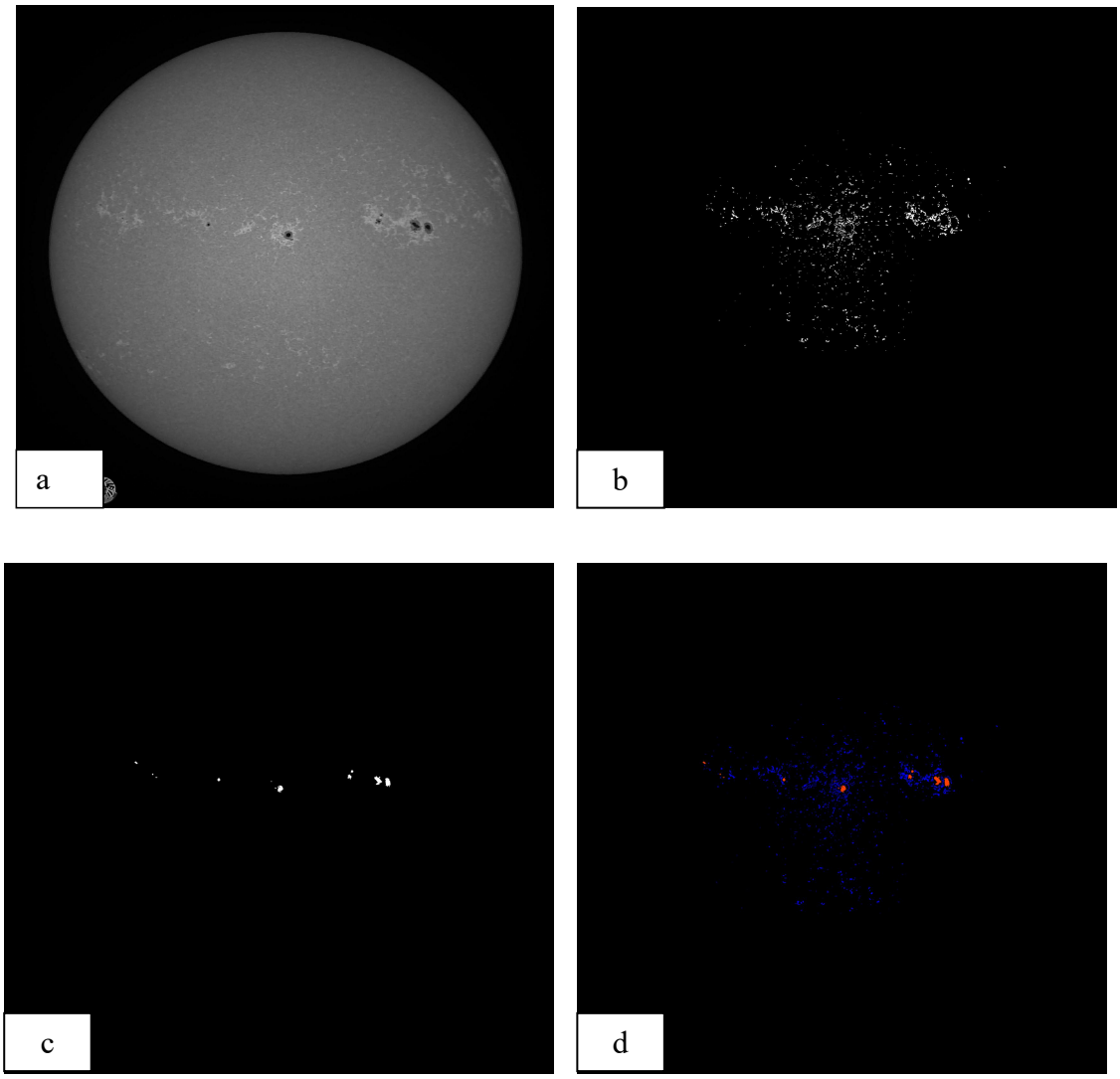


Figure 96: Image at 393 nm on 1st October 2011 at time 02:47: (a) Original image; (b) Bright regions segmentation; (c) Sunspot segmentation; (d) FF for bright regions in blue and sunspot in orange

Table 15 shows the FF of the wavelength 393 nm images taken on 1st October 2011 at 02:47; this method has been applied to all available images to calculate their FF. The FF results still need verification, and verification requires another resource of information. Milbourne et al. (Milbourne et al.,

2019) calculated the FF but in a different time period and the images taken by a very high-resolution telescope were not valid for SODISM.

Table 15: Example of numbers of bright regions in the different parts of the solar disc in images taken 1st October 2011 at 02:47

Index	RI	RO	Strip area/pixel	Bright region area/pixel
1	0	0.07	1.2×10^4	5×10^2
2	0.07	0.16	5.2×10^4	26.7×10^2
3	0.16	0.25	9.3×10^4	26.8×10^2
4	0.25	0.35	1.5×10^5	37×10^2
5	0.35	0.45	2.0×10^5	34×10^2
6	0.45	0.55	2.5×10^5	15×10^2
7	0.55	0.65	3.0×10^5	35×10^2
8	0.65	0.75	3.5×10^5	0.01×10^2
9	0.75	0.85	4.0×10^5	0.25×10^2
10	0.85	0.95	4.5×10^5	0.00
11	0.95	1	2.4×10^5	0.00
Total of bright regions				14931

4.3 Filling factor catalogue

In the literature, the FF, specified as a function of radial position on the solar disc, is computed for each image and recorded in a tabular manner. However, this does not provide extensive information about the experiments conducted on tens or hundreds of such images. To provide a means of summarising these results so that a single glance at a table would give an overall picture of the work done, a cataloguing procedure was proposed. To the best of the author's knowledge, this is the first time that such a FF catalogue has been developed for SODISM images. Parametric and non-parametric statistics (Sheskin 2003) are used to describe the data. While parametric statistics make the assumption that the data is drawn from a population with a normal distribution, non-parametric tests make no such assumption. Hence, working with both will provide a robust correction to counter the disadvantages of each type of statistical test; thus, the results should complement each other.

The parametric measures employed are the three statistical measures of mean, variance and skewness. In comparison, medians and inter-quartile ranges (IQRs) are used as non-parametric measures. Since skewness gives information about the normality of the distribution, it could be expected that the results of the first and second statistical measures would be reliable if the third (higher) measure indicates normality. On the other hand, an individual may rely on information provided by the non-parametric statistics if the third order measure indicates non-normality. As a rule of thumb, researchers assume that a value of skewness above ±2 implies a distribution that is highly skewed and hence not normally distributed (Field 2013; Gravetter and Wallnau 2016; Trochim and Donnelly 2001). Furthermore, the IQR can be used to compute the number of outliers in the data (Rousseeuw and Hubert 2011) using the following equation:

$$for\ x\ \in observations \quad (14)$$

where x is an outlier if it satisfies any of the two conditions.

$$x \begin{cases} < q_1 - 1.50\ IQR \\ > q_3 + 1.5\ IQR \end{cases}$$

where q_1 and q_3 are the first and third quartiles respectively.

Thus, the proposed cataloguing table will contain the information listed below;

- Mean-Total Sunspots μ.
- Variance-Total Sunspots σ (standard division).
- Skewness-Total Sunspots
- Median-Total Sunspots
- Interquartile-range-Total Sunspots
- Total Sunspot Outliers

The above variables have been calculated for whole FF SODISM images and those results are compared with the ratio FF SODISM results. This comparison is made to confirm FF calculations provide correct results in a normal distribution and matching between FF results and FF ratio results.

- Mean-Sunspot Ratio
- Variance-Sunspot Ratio
- Skewness-Sunspot Ratio
- Median-Sunspot Ratio
- Interquartile-range- Sunspot Ratio
- Sunspot Outliers Ratio

Table 16: Proposed cataloguing table

Total number of samples $n = 206$	
Test	**Result**
μ (Total Sunspots)	1.87×10^{3}
σ (Total Sunspots)	1.65×10^{3}
Skewness (Total Sunspots)	111×10^{-2}
Med (Total Sunspots)	1.57×10^{3}
IQR (Total Sunspots)	2.1×10^{3}
Total Sunspot Outliers	7
μ (Sunspot Ratio)	6.6×10^{-4}
σ (Sunspot Ratio)	5.8×10^{-4}
Skewness (Sunspot Ratio)	113×10^{-2}
Med (Sunspot Ratio)	5.5×10^{-4}
IQR Sunspot Ratio)	7.5×10^{-4}
Sunspot Outliers Ratio	7

From the results presented in Table 16, the skewness measure indicates that the data are normally distributed and thus, the parametric scores are reliable. It is therefore possible to report that the extensive evaluation conducted on 206 SODISM data images achieved a mean sunspot region of 1.87×10^{3} in total, and a corresponding average sunspot ratio of 6.6×10^{-4}. It can further be deduced that any image picked from the dataset is expected to have a FF signature of ±3 standard deviation; moreover, the value of σ can be conveniently retrieved from Table 16. Furthermore, the IQR-based outlier detection has revealed that there are seven images within the data that lie outside the normal distribution. Hence, as with any other real-world data, it is expected that these few outliers ($\approx 0.033\%$ of the dataset) will exhibit some irregularities. The FF catalogue

holds records of SODISM images information; this catalogue will organize as follows:

Reading from left to right, the first column records the sunspot date in the dd/mm/yy format, the second column indicates the time of sunspot in SODISM, and the third is the equivalent time in NOAA catalogue. The fourth and fifth columns show the number of sunspots in SODISM images and the equivalent number of sunspots in the NOAA catalogue. The location of sunspots captured by SODISM is recorded in the sixth column, and the size of sunspots in SODISM catalogue is shown in column seven. Finally, the filling factor ratio shows in the final eighth column. Figure 97 shows an example for the FF catalogue.

Date	Time of Sunspot in SODSIM	Time of NOAA Sunspot	number of sunpots SODISM	number of spots NOAA	Loction of Sunspot SODISM	Size of Sunspots SODISM	Total FF Ratio
05/08/2010	04:49	07:35	4	4	N13W24	530	0.000531
05/08/2010	04:49	07:35	4	4	N11E67		
05/08/2010	04:49	07:35	4	4	S18E56		
05/08/2010	04:49	07:35	4	4	N25W49		

Figure 97: Filling factor catalogue for SODISM images

The data is available from August 2010 to January 2014, and that explains that figures 98 shows only one month in 2014.

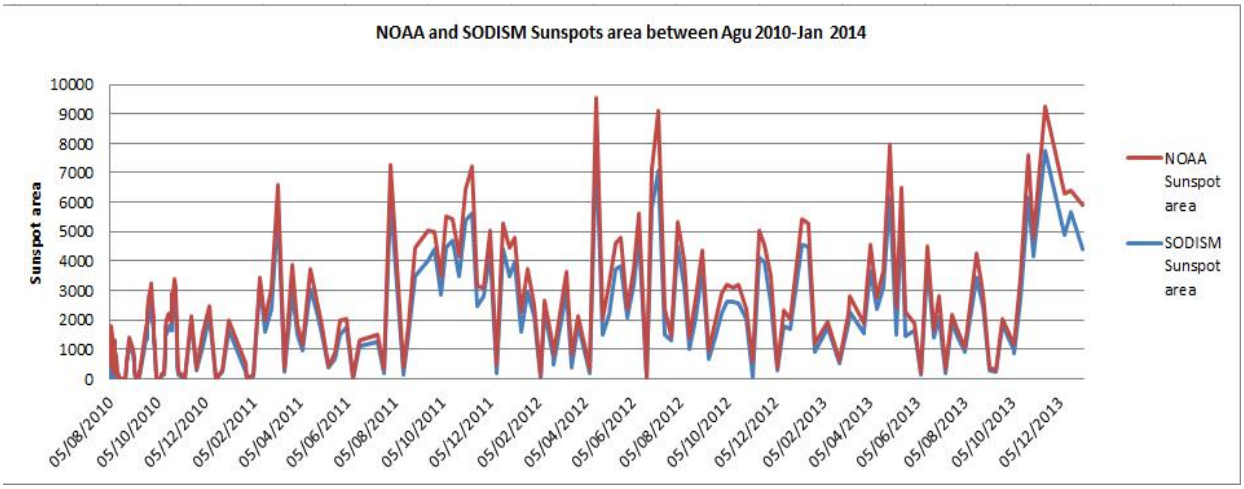


Figure 98: Sunspot areas extracted from SODISM by the automatic method and sunspot areas provided by NOAA in Aug. 2010-Jan 2014

The daily sunspot areas illustrated in figure 98 extracted by the automatic method from SODISM images are compared with those available as TXT files at NOAA catalogue. The horizontal axis represents the date of image capture, the vertical the sunspot area, in units of millionths of a solar hemisphere. The correlation coefficient in the same tendency is 93% shown in figure 99.

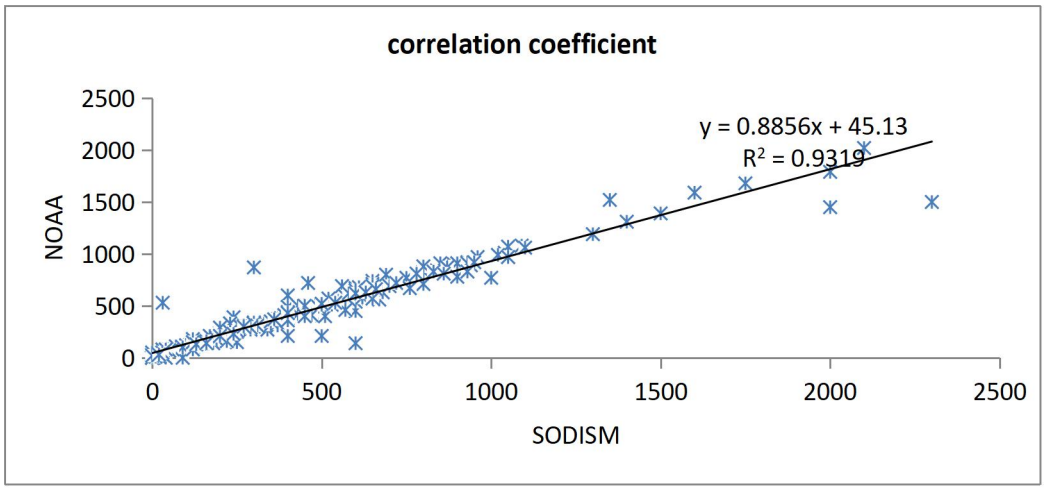


Figure 99: The correlation coefficient between NOAA and SODISM sunspot areas

This high accuracy of recognition appears to indicate that the method works well.

The discrepancy in figure between the SODISM and NOAA catalogues may be caused by comparing of SODISM and NOAA images on the same days without accounting for varying collection time. Moreover, the effect caused by instrument noises indistinguishable from smaller sunspots leads to sunspot false identification.

4.4 Conclusion

This chapter has shown that the SODISM FF catalogue created in this research is a novel means of summarizing information and details of available SODISM images including size, number and location of sunspots.

The comparison has been made between SODISM catalogue and NOAA catalogue relating to a number of sunspots detected in the same time period. Their location, size and the accuracy of comparison for the period August 2010 to January 2014 shows a correlation coefficient of 95.17%. Moreover, the sizes of sunspots are matching 93% in NOAA catalogue.

In summary, an innovative statistical means has been presented to characterise SODISM data. The afore-defined metrics present researchers with a new method of summarising hundreds of experiments in a concise and easily understood manner. The catalogue can then be used as a template to conduct comparisons. Should researchers embark on a study similar to this thesis, it is hoped that they will use this novel procedure to present their results as a means for comparison

A significant practical contribution to a field of knowledge has also been provided in that a cataloguing procedure has been proposed and implemented, which gives a clear representation of the SODISM data. It is hoped that the proposed technique will assist researchers in their future work.

CHAPTER FIVE

5. System evaluation and assessment

5.1 Assessment of image quality enhancement algorithm

Two approaches to evaluate the performance of the proposed image quality enhancement algorithm have been adopted; visual observation and a metric-based method. Observation of the visual representations has already been described in Chapter 4 (Section 4.2). Here, a mathematical metric approach is used to assess whether significant improvement has been achieved. The mean square error (MSE) was used as a measure of accuracy, and was then compared with the bicubic interpolation method. Interestingly, the results show a clear improvement for the maps of the five wavelengths, and the MSE calculated for the proposed technique is less than that of the bicubic method. Table 17 shows the results for a sample image. Furthermore, Figure 100 illustrates the detailed results for all wavelengths. As a second metric for comparison, the peak signal-to-noise ratio (PSNR) was computed. An example of the results for images taken on 1st October 2011 is shown in Table 18, and once again the proposed technique proved superior to both the bicubic and WZP techniques. The superiority of the proposed technique is clear again in Figure 101, where, during the lifetime of SODISM, the enhancement technique showed better results in resolution than other methods, both visual assessment and quantitative calculation assessment, confirming the efficacy of the proposed enhanced method. Additionally, Structural Similarity Index Metric (SSIM) has been calculated in the proposal method and bicubic interpolation technique, and an example of the results is shown in table 16 for five wavelengths.

Table 17: Comparison of MSE and SSIM values obtained by bicubic interpolation and using the proposed technique for images at different wavelengths taken on 1st October 2011

Wavelength and a time at 1st October 2011	MSE Bicubic	MSE Proposed Tec.	SSIM Bicubic	SSIM Proposed Tec.
215 nm at 02:55	19.6×10^{-3}	5×10^{-3}	595×10^{-3}	785×10^{-3}
393 nm at 02:47	96.6×10^{-3}	24×10^{-3}	654×10^{-3}	800×10^{-3}
535 nm at 12:30	191×10^{-3}	48.9×10^{-3}	690×10^{-3}	828×10^{-3}
607 nm at 01:01	353×10^{-3}	89.6×10^{-3}	720×10^{-3}	848×10^{-3}
782 nm at 02:07	308×10^{-3}	79×10^{-3}	701×10^{-3}	802×10^{-3}

The proposed method works well with all five wavelengths, despite the fact that images at each wavelength are of varying quality and resolution, and the results show a consistent improvement even with the poor quality of images taken at a wavelength of 215 nm.

The Structural Similarity Index Metric (SSIM) is more accurate and more consistent than MSE and PSNR.SSIM is also high performance (Varnan et al., 2011), the greater values shows greater image similarity, in table 17 shows that proposal method values of SSIM is greater than bicubic method.

Table 18: Comparison of PSNR values obtained by bicubic interpolation and the proposed technique at different wavelengths

Wavelength and time 1st October 2011	Peak signal-to-noise ratio (PSNR)		
	Bicubic interpolation	WZP	Proposed technique
215 nm at 02:55	4115×10^{-2}	4703×10^{-2}	4710×10^{-2}
393 nm at 02:47	3421×10^{-2}	4021×10^{-2}	4024×10^{-2}
535 nm at 12:30	3124×10^{-2}	3522×10^{-2}	3717×10^{-2}
607 nm at 01:01	2859×10^{-2}	3452×10^{-2}	3454×10^{-2}
782 nm at 02:07	2917×10^{-2}	3504×10^{-2}	3509×10^{-2}

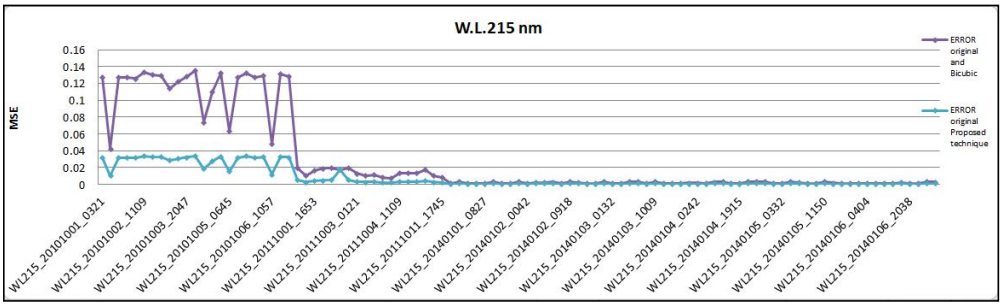
Figure 101 shows that the PSNR for the proposed technique has greater values than the WZP and bicubic techniques; thus, the results confirm

the superiority of the proposed method. To support the enhancement results of the proposal method, NAE, SC and NCC have been calculated. Table 19 shows the normalized absolute error is very low for all wavelengths, and the NCC is approximately (0.96) which reflects excellent correlation between the original and enhancement images, and low values of Structural Content shows its good quality images (Memon, Unar and Memon, 2015) while the bicubic method shows high NAE, which indicates the proposal enhancement method achieved excellent results compared with the bicubic interpolation method. Moreover the high Structural Content in bicubic interpolation indicates a poor image (Memon, Unar and Memon, 2015).

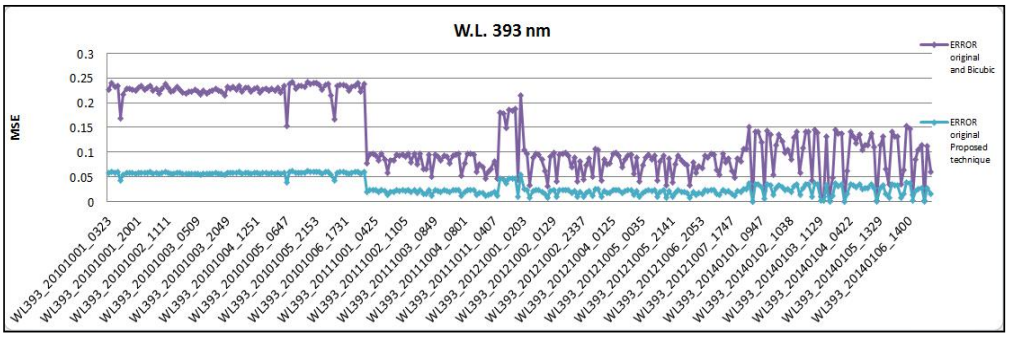
Table 19: Comparison of NEA, NCC and SC values obtained by bicubic interpolation and using the proposed technique for images at different wavelengths taken on 1st October 2011

Wavelength and a time at 1st October 2011	Bicubic interpolation			Proposed technique		
	NAE	NCC	SC	NAE	NCC	SC
215 nm at 02:55	23.02×10^{-2}	641×10^{-3}	1962×10^{-3}	25.9×10^{-3}	960×10^{-3}	1002×10^{-3}
393 nm at 02:47	23.01×10^{-2}	608×10^{-3}	1958×10^{-3}	25.9×10^{-3}	959×10^{-3}	1002×10^{-3}
535 nm at 12:30	23.01×10^{-2}	609×10^{-3}	1958×10^{-3}	25.9×10^{-3}	959×10^{-3}	991×10^{-3}
607 nm at 01:01	23.01×10^{-2}	609×10^{-3}	1958×10^{-3}	25.9×10^{-3}	959×10^{-3}	992×10^{-3}
782 nm at 02:07	23.03×10^{-2}	608×10^{-3}	1958×10^{-3}	25.9×10^{-3}	959×10^{-3}	1000×10^{-3}

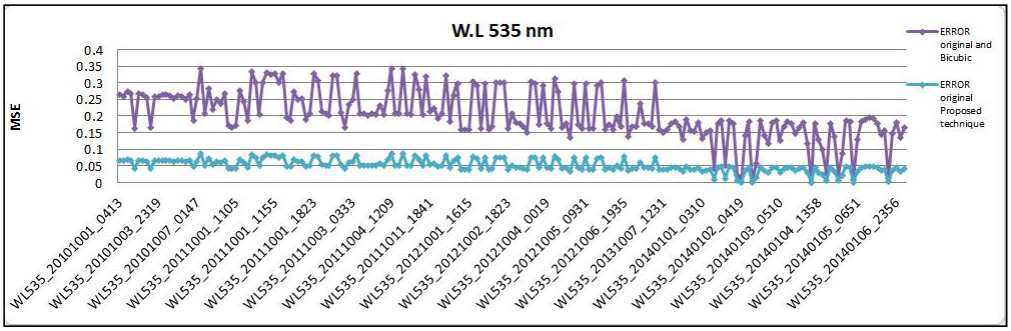
Additional calculations have been completed for all data set to compare Normalized absolute error (NAE), Normalized Cross-Correlation (NCC) and Structural Content (SC). The results set out on table 19 support the previous assumptions, where the NCC >0.9 on all images, and this indicates a good match (Rao, Prathapani and Nagabhooshanam, 2014) between the benchmark and enhancement images. Furthermore, NAE is less than 26×10^{-3} in all enhancement images, and this small value reflect excellent enhancement image results (Khidse, 2014). Finally, measurement of image quality after enhancement is very important to numerous image processing applications. The mean-squared-error (MSE) is the simplest, and the most widely used. Particularly, SSIM has been added to support the evaluation method.



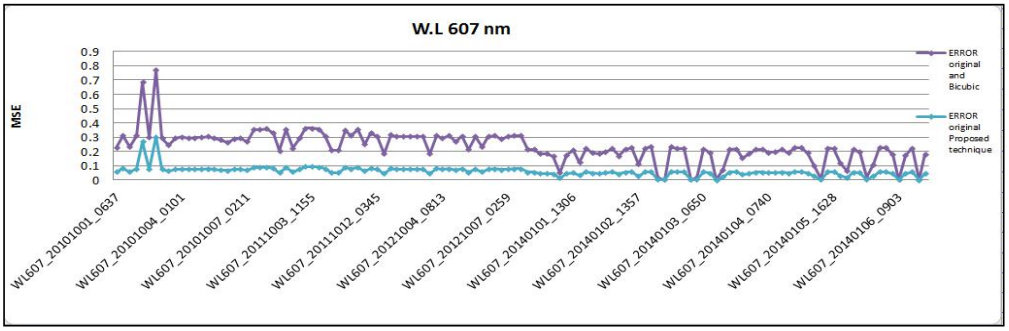
a



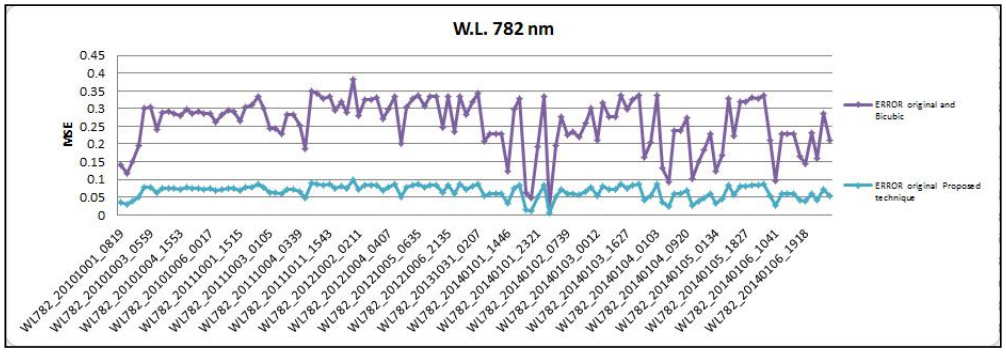
b



c

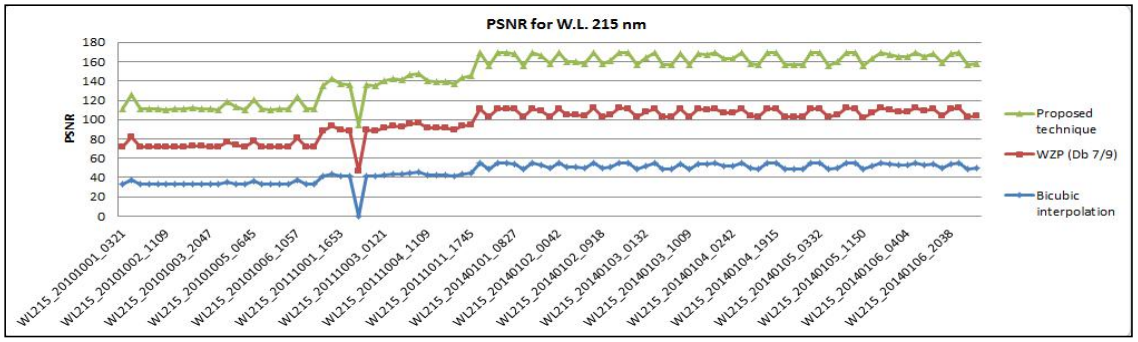


d

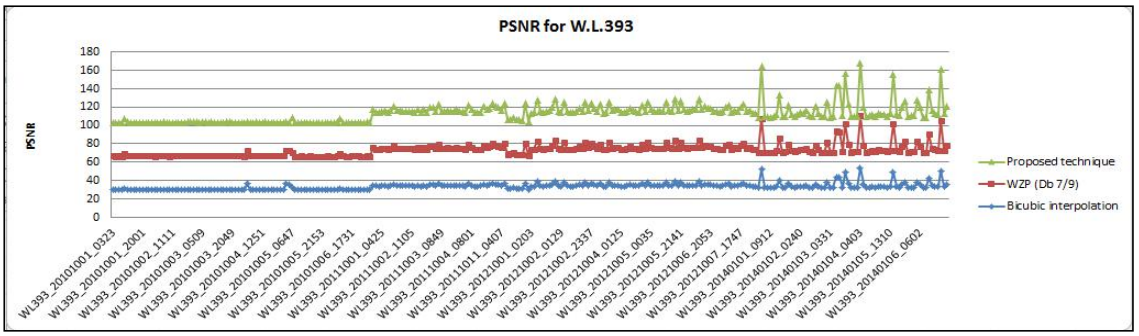


e

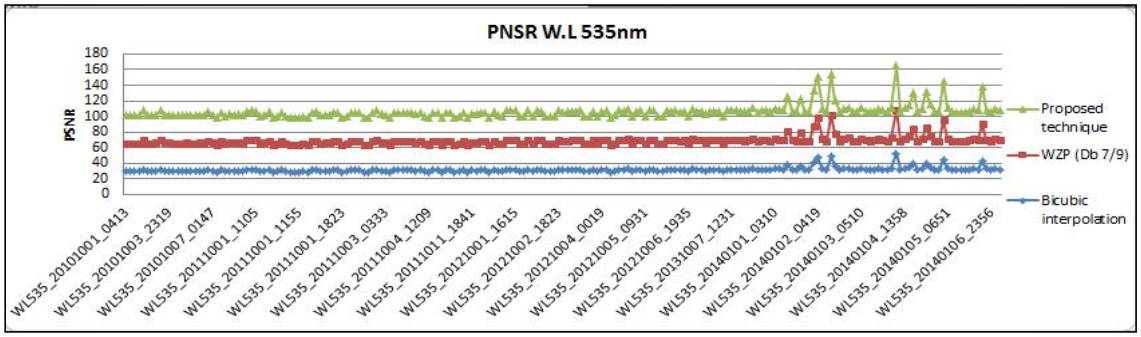
Figure 100: Comparison of MSE values obtained by bicubic interpolation and the proposed technique for images at 5 wavelengths taken over the period 2010-2014: (a) 215 nm, (b) 393 nm, (c) 535 nm, (d) 607 nm, (e) 782 nm



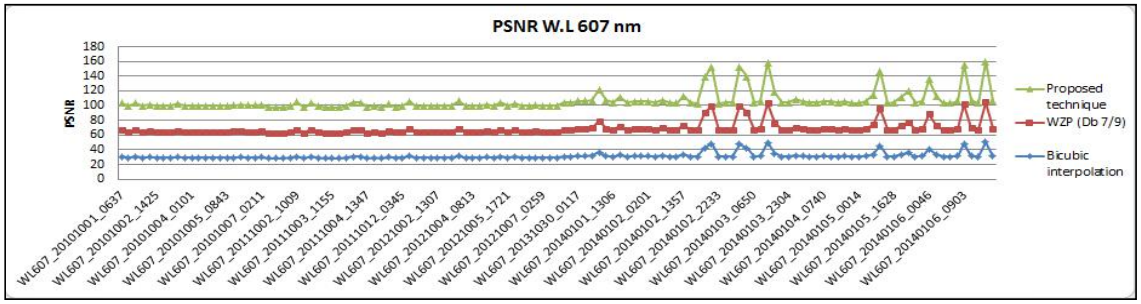
a



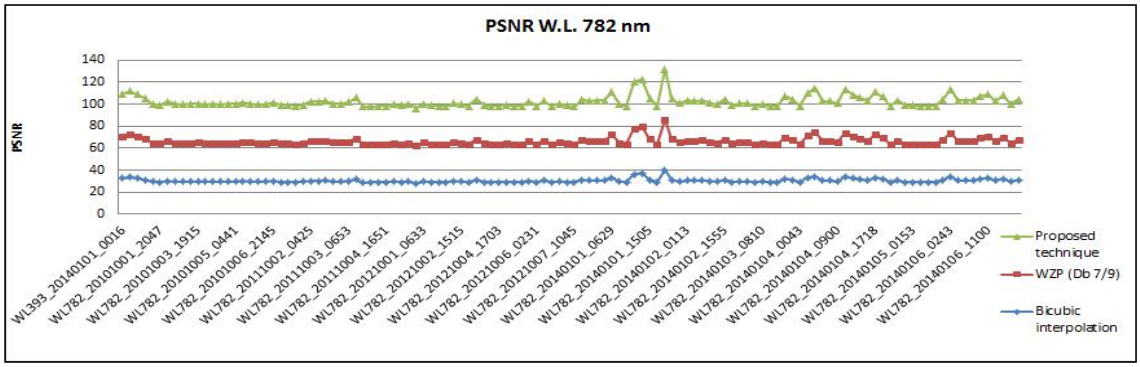
b



c



d



e

Figure 101: Comparison of PSNR values obtained by bicubic interpolation, WZP and the proposed technique for images at 5 wavelengths taken over the period 2010-2014: (a) 215 nm, (b) 393 nm, (c) 535 nm, (d) 607 nm, (e) 782 nm

Figure 102 shows that the average value of NAE for the proposed technique has lower values than the bicubic technique; thus, the results confirm the superiority of the proposed method. To support the proposal method of enhancement results, the average values of the NCC have been shown in figure 103, and the values of NCC for the proposed method are greater than the bicubic interpolation method, therefore new evaluations show that the proposed enhancement method achieved its goal with excellent results. .

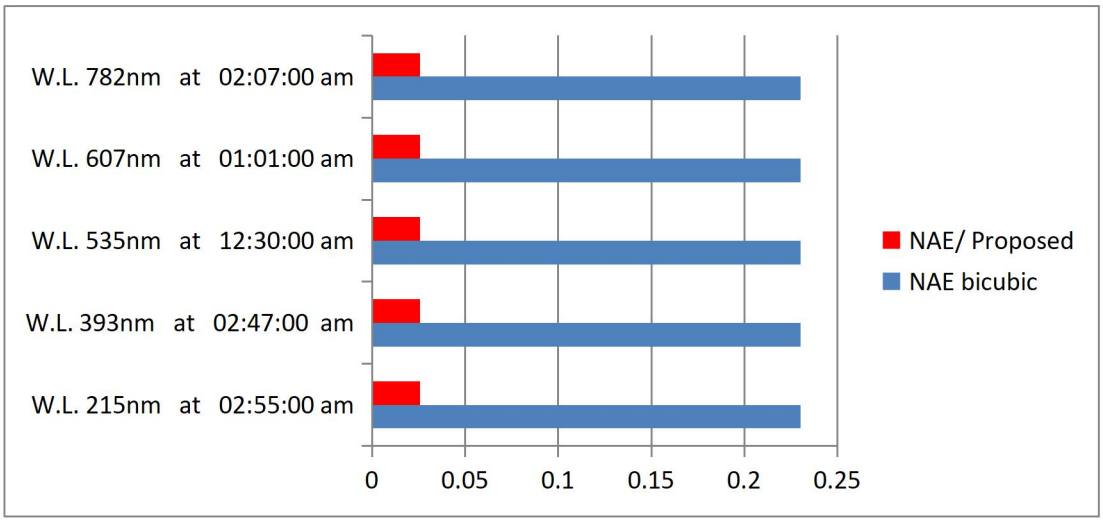


Figure 102: Comparison of average NAE values obtained by bicubic interpolation, and the proposed technique for images at 5 wavelengths taken over the period 2010-2014

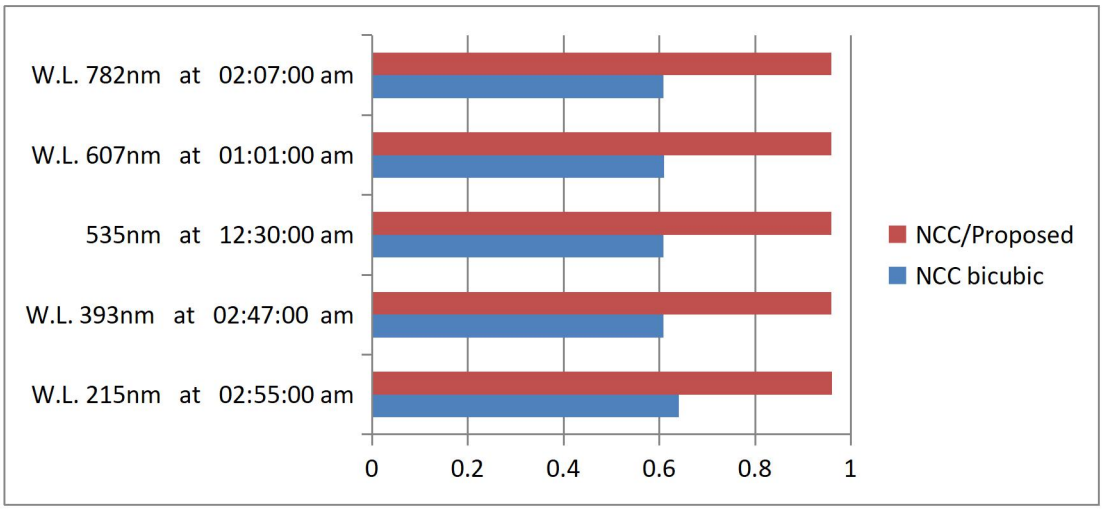


Figure 103: Comparison of average NCC values obtained by bicubic interpolation, and the proposed technique for images at 5 wavelengths taken over the period 2010-2014

5.2 Assessment of segmentation algorithm

Once again, two approaches were used to evaluate the segmentation performance: firstly, visual observation was used to compare the output of the segmentation algorithm with the ground truth results. Samples of the segmentation output are shown in Figures 104-106.

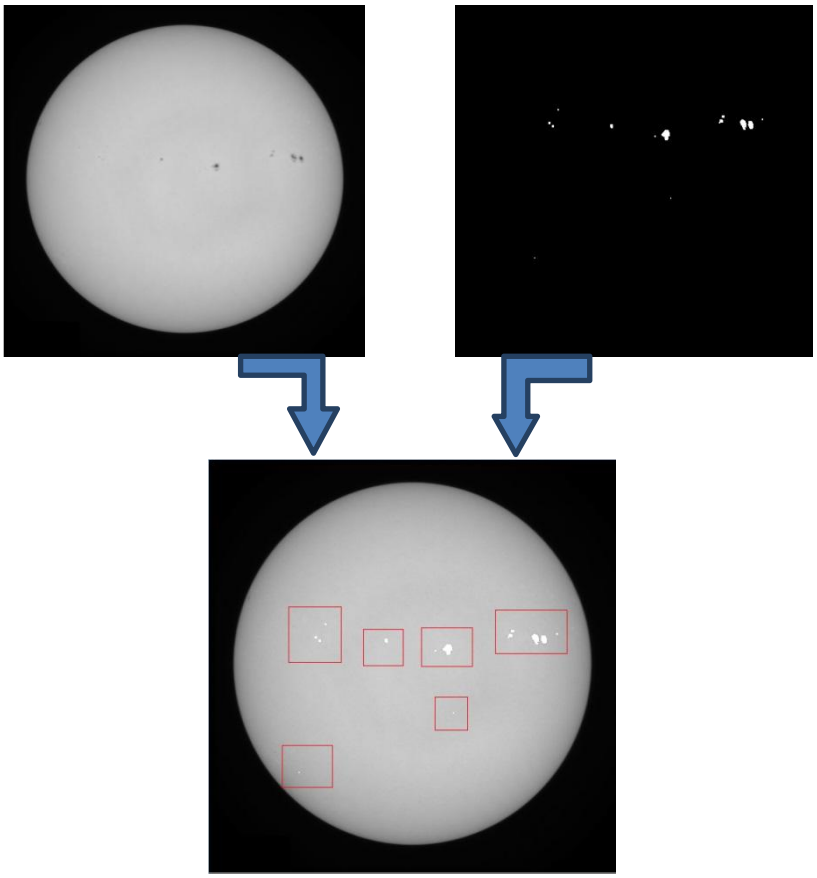


Figure 104: Visual matching of sunspot position after detection for image at 782 nm taken on 1st October 2011

Figure 104 shows that sunspots are matched in the same location and are of the same size. Therefore, visual assessment shows perfect matches. In Figure 105, the images at the 607 nm wavelength illustrate the visual assessment, and Figure 106 shows an example of a 535 nm wavelength image where a visual match was made between the original and the sunspot image.

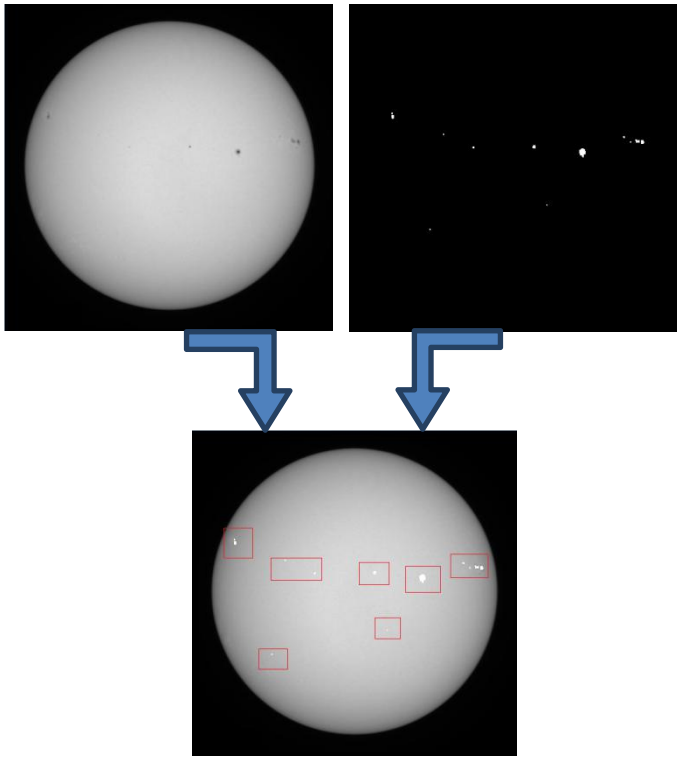


Figure 105: Figure 105: Visual matching of sunspot position after detection for image at 607 nm taken on 3rd November 2011

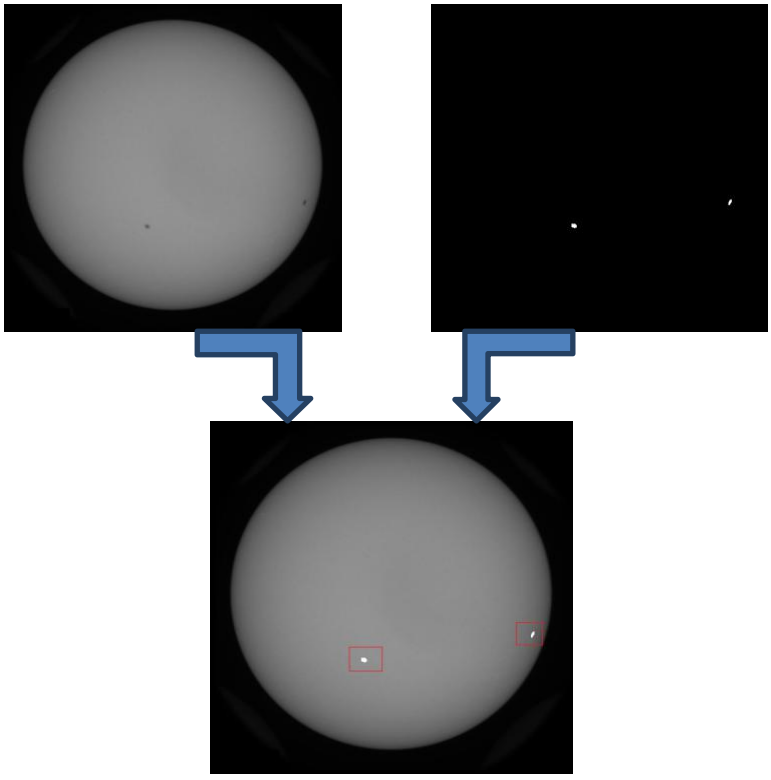


Figure 106: Visual matching of sunspot position after detection for image at 535 nm taken on 7th October 2012

Finally, Figure 107 shows the results of visual assessment applied to a 393 nm wavelength image, which illustrates that sunspots are matched at the same location as the original image.

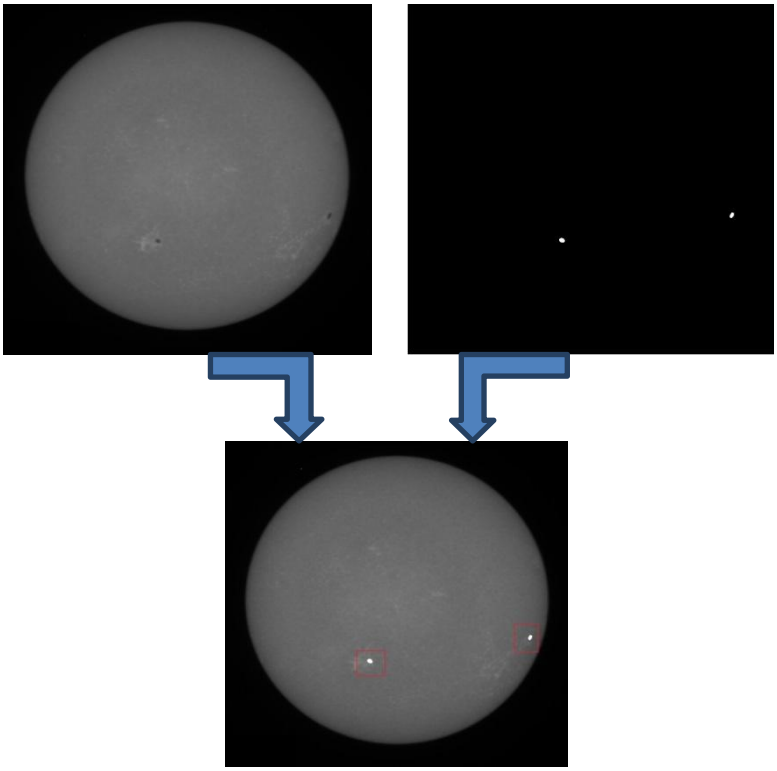


Figure 107: Visual matching of sunspot position after detection for image at 393 nm taken on 7th October 2012

5.3 Filling factor assessment

Obvious improvements were observed in the FF calculated at 535 nm and 393 nm wavelength images before and after applying the proposed enhancement techniques. The calculations presented in Tables 20 and 21 show the results for the enhanced images with values nearer to those of the benchmark images, and Figure 108 shows the details of the comparison.

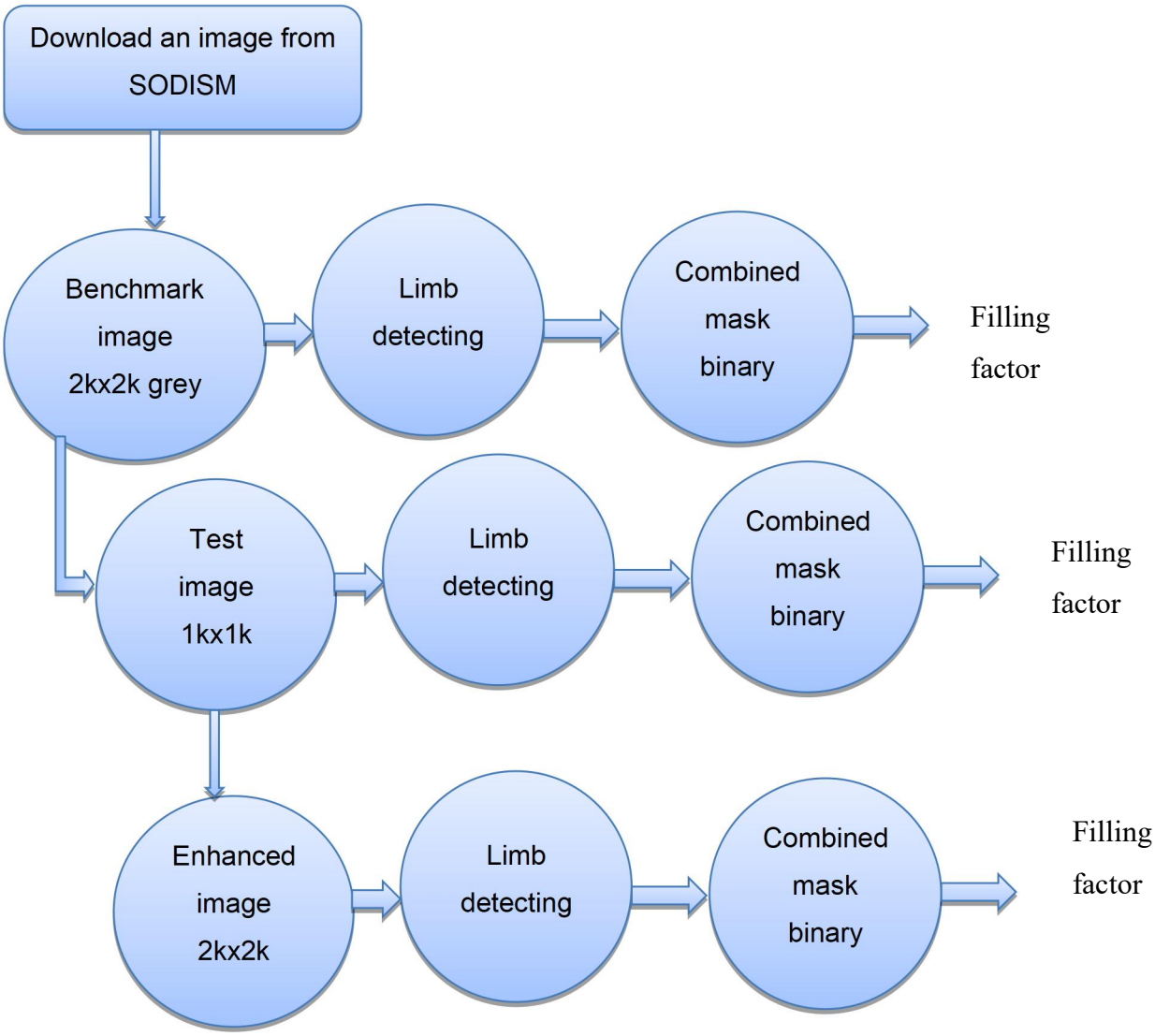


Figure 108: Compression procedure for the FF results

Table 20 shows the FFs for the benchmark images taken at a 393 nm wavelength on 1st October 2011 with a total of 4107 pixels.

Table 20: FF for the benchmark images at 393 nm, on 1st October 2011

Idx	RI	RO	Strip area/ pixel	Sunspot area pixels
1	0	0.07	1.2×10^{4}	54607×10^{-2}
2	0.07	0.16	5.2×10^{4}	56390×10^{-2}
3	0.16	0.25	9.3×10^{4}	6403×10^{-2}
4	0.25	0.35	1.5×10^{5}	17000×10^{-2}
5	0.35	0.45	2.0×10^{5}	57800×10^{-2}
6	0.45	0.55	2.5×10^{5}	21484×10^{-2}
7	0.55	0.65	3.0×10^{5}	183416×10^{-2}
8	0.65	0.75	3.5×10^{5}	5500×10^{-2}
9	0.75	0.85	4.0×10^{5}	5100×10^{-2}
10	0.85	0.95	4.5×10^{5}	0
11	0.95	1	2.4×10^{5}	3000×10^{-2}
Number of feature pixels:				4107

Table 21 shows the FFs for an enhanced image at 393 nm taken on 1st October 2011 with a total of 4132 pixels, and Table 21 shows the FFs for test images at the same wavelength and with a total of 900 pixels.

Table 21: FF for enhanced image at 393 nm, taken on 1st October 2011

Idx	RI	RO	Strip area/pixel	Sunspot area pixel
1	0	0.07	1.2×10^{4}	13544×10^{-2}
2	0.07	0.16	5.2×10^{4}	82456×10^{-2}
3	0.16	0.25	9.3×10^{4}	5600×10^{-2}
4	0.25	0.35	1.5×10^{5}	10375×10^{-2}
5	0.35	0.45	2.0×10^{5}	62737×10^{-2}
6	0.45	0.55	2.5×10^{5}	6220×10^{-2}
7	0.55	0.65	3.0×10^{5}	211069×10^{-2}
8	0.65	0.75	3.5×10^{5}	6800×10^{-2}
9	0.75	0.85	4.0×10^{5}	3600×10^{-2}
10	0.85	0.95	4.5×10^{5}	2800×10^{-2}
11	0.95	1	2.4×10^{5}	8000×10^{-2}
Check column sums:				4132

Table 22: FF for a test image at 393 nm, taken on 1st October 2011

Idx	RI	RO	Strip area/ pixel	Sunspot area pixel
1	0	0.07	3.1×10^{3}	12773×10^{-2}
2	0.07	0.16	1.3×10^{4}	8727×10^{-2}
3	0.16	0.25	2.3×10^{4}	0
4	0.25	0.35	3.8×10^{4}	4500×10^{-2}
5	0.35	0.45	5.0×10^{4}	9600×10^{-2}
6	0.45	0.55	6.3×10^{4}	30000×10^{-2}
7	0.55	0.65	7.6×10^{4}	23700×10^{-2}
8	0.65	0.75	8.8×10^{4}	400×10^{-2}
9	0.75	0.85	10.1×10^{4}	0
10	0.85	0.95	11.4×10^{4}	300×10^{-2}
11	0.95	1	6.1×10^{4}	0
Number of feature pixels:				900

The percentage between the benchmark and test image from one side, and the benchmark and enhancement image from the other side shows that the enhanced image is closer to the benchmark image, as seen in Table 23. This comparison is made using the following equation;

$$different\ \% = \left|\frac{benchmark - test\ or\ enhanched}{benchmark}\right| \quad (14)$$

Table 23 shows that the enhanced images (seven cases) are close to the benchmark, and the total number of pixels from the enhanced method almost matches the benchmark with a difference of only 0.6%.

The results in table 23 used "✓" to indicate that the enhanced images are closer to the benchmark images than the test images and "×" for the opposite case, while "0" is entered if there is no value.

Table 23: Comparison of the enhanced and test images with the benchmark image.

Benchmark to Enhanced image	Benchmark to Test image	Result
7520×10^{-2}	644×10^{-2}	X
4623×10^{-2}	3809×10^{-2}	X
1254×10^{-2}	10000×10^{-2}	✓
3897×10^{-2}	588×10^{-2}	X
854×10^{-2}	3356×10^{-2}	✓
7105×10^{-2}	45856×10^{-2}	✓
1508×10^{-2}	4831×10^{-2}	✓
2364×10^{-2}	7091×10^{-2}	✓
2941×10^{-2}	10000×10^{-2}	✓
0	0	0
16667×10^{-2}	10000×10^{-2}	X
60.8×10^{-2}	1235×10^{-2}	✓

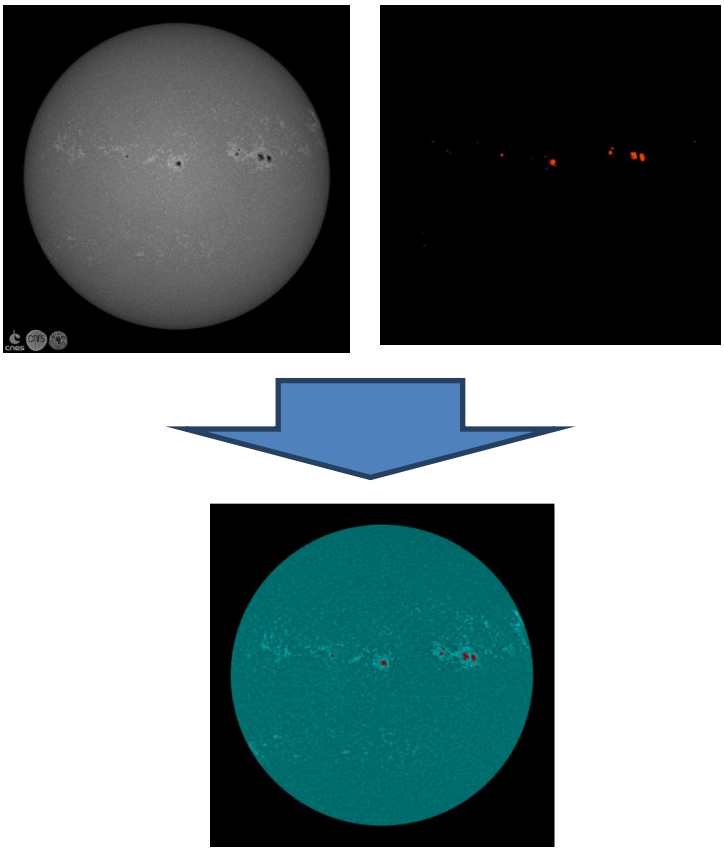


Figure 109: Visual comparison shows matching sunspots with the benchmark image in a 393 nm example

Figure 109 also shows the results for images at 393 nm, where the detected sunspot and the benchmark image closely match. Repeating the same comparison with data at the 535nm wavelength gave the results presented in Table 23 for benchmark images. FFs for enhanced images in table 24, and FFs for test images in table 25, have been captured on the same date to make a comparison between the enhancement images and test images. The results in table 26 show six cases where the enhanced images are closest to the benchmark image.

Table 24: FFs for benchmark images at 535 nm on 1st October 2011

Idx	RI	RO	Strip area/ pixel	Sunspot area pixel
1	0	0.07	1.2×10^4	0
2	0.07	0.16	5.2×10^4	1.1×10^3
3	0.16	0.25	9.3×10^4	0
4	0.25	0.35	1.5×10^5	0.1×10^3
5	0.35	0.45	2.0×10^5	0
6	0.45	0.55	2.5×10^5	0.5×10^3
7	0.55	0.65	3.0×10^5	1.2×10^3
8	0.65	0.75	3.5×10^5	0.8×10^3
9	0.75	0.85	4.0×10^5	0.6×10^3
10	0.85	0.95	4.5×10^5	0
11	0.95	1	2.4×10^5	0
Check column sums.				3975

Table 25: FFs for enhanced image at 535 nm taken on 1st October 2011

Idx	RI	RO	Strip area/ pixel	Sunspot area pixel
1	0	0.07	1.2×10^4	0
2	0.07	0.16	5.2×10^4	1.1×10^3
3	0.16	0.25	9.3×10^4	0
4	0.25	0.35	1.5×10^5	0.1×10^3
5	0.35	0.45	2.0×10^5	0
6	0.45	0.55	2.5×10^5	0.4×10^3
7	0.55	0.65	3.0×10^5	1.1×10^3
8	0.65	0.75	3.5×10^5	0.9×10^3
9	0.75	0.85	4.0×10^5	0.4×10^2
10	0.85	0.95	4.5×10^5	0
11	0.95	1	2.4×10^5	0
Check column sums				4052

Table 26: FFs for test image at 535 nm taken on 1st October 2011

Idx	RI	RO	Strip area/ pixel	Sunspot area pixel
1	0	0.07	3.1×10^3	0
2	0.07	0.16	1.3×10^4	0.3×10^3
3	0.16	0.25	2.3×10^4	0
4	0.25	0.35	3.8×10^4	0.02×10^3
5	0.35	0.45	5.0×10^4	0
6	0.45	0.55	6.3×10^4	0.1×10^3
7	0.55	0.65	7.6×10^4	0.3×10^3
8	0.65	0.75	8.8×10^4	0.3×10^3
9	0.75	0.85	10.1×10^4	0.01×10^3
10	0.85	0.95	11.4×10^4	0
11	0.95	1	6.1×10^4	0
Check column sums:				1114

Table 26, as previously mentioned, shows the comparison between the results of the benchmark, the enhanced and the test images. The results show that enhanced images match the benchmark images more closely than the test images, again supporting visual and computational results as previously stated.

Again as in table 23, the results in table 27 use “ ✓ ” to indicate if the enhanced images are closer to the benchmark than the test images, and “×” for the opposite case, while "0" indicates there is no value.

Table 27: Comparison of an enhanced and test image to benchmark image at 535 nm

Benchmark to Enhanced image	Benchmark to Test image	Result
0	0	0
576×10^{-2}	1171×10^{-2}	✓
0	0	0
368×10^{-2}	4105×10^{-2}	✓
0	0	0
259×10^{-2}	1777×10^{-2}	✓
149×10^{-2}	782×10^{-2}	✓
548×10^{-2}	4804×10^{-2}	✓
2000×10^{-2}	1333×10^{-2}	X
0	0	0
0	0	0
194×10^{-2}	1210×10^{-2}	✓

Finally, more assessments have been made by visual comparison and figure 110 shows a random example for W.L 535 nm shows the results of a sunspot in benchmark image is closer to the improved image than the test image.

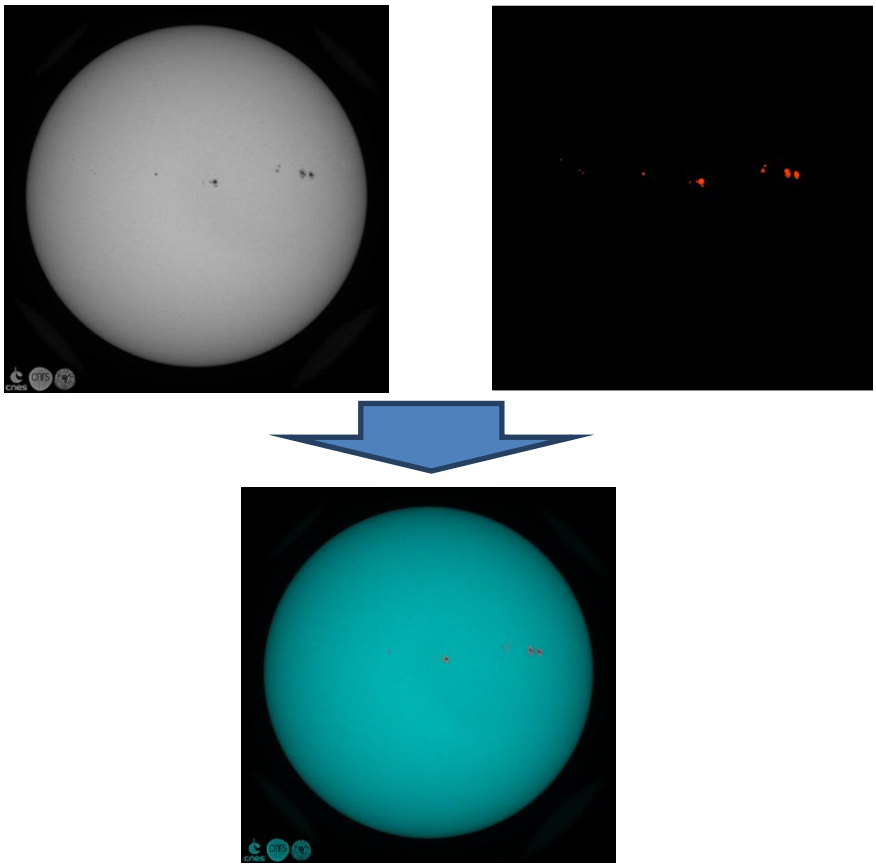


Figure 110: Visual comparison of sunspots with the benchmark image for an example at 535 nm

5.4 Summary

According to the peak signal-to-noise ratio (PSNR) and the mean square error (MSE) results in Figures 100 and 101, the results of the proposed method indicate a better performance compared with those of the bicubic methods. Additionally, figure 102 and 103 for NEA and NCC shows the superiority of the proposed method, and the values of SSIM in table 17 show that the proposed techniques have delivered a good performance Furthermore, the results displayed in Tables 23 and 27 concerning the enhancement method, which was applied to images of 393 nm and 535 nm wavelengths ,show that the values of the enhanced images are closer to

those of the benchmark than the test images, reflecting the improvement in the enhanced images. By choosing random images to show the matching of segmentation results with benchmark images (as shown in Figures 104-106), another comparison has been conducted in this chapter. All results from the NAE, NCC, CS, PSNR, MSE, SSIM and FF calculations, besides the visual comparisons, indicate that the proposed methods have been successful in achieving the aim of this study. In this chapter, different performance measures were used to evaluate the proposed algorithms for image quality enhancement and segmentation. Furthermore, the FF calculations have been used to compare SODISM and SOHO data. In all of the assessments, the proposed algorithms demonstrated excellent performance in comparison with existing techniques. Additionally, SSIM have been calculated, and the average for the proposal method was 0.82, very close to the unit 1 which represents the original image values. It is also better than the values of the bicubic method, and the NAE of proposal method is <0.025, which reflects good results, and NCC shows high image qualities in all wavelengths.

CHAPTER SIX

6. Conclusion and future work

6.1 Overall conclusion

The main achievement of the research described in this thesis is the design of a collection of algorithms related to the automatic detection of solar sunspots from SODISM images, which have been developed from multiple technologies. Furthermore, the effectiveness of the use of a novel method for the enhancement of SODISM images taken at different wavelengths has been demonstrated. A further achievement is the development of a FF catalogue for SODISM images for use by researchers. This outcome is unique because previously no reliable FF catalogue for SODISM images has been available for use in conducting scientific comparisons between different wavelengths.

6.2 Detailed conclusions

The goals and specific achievements of this thesis are outlined as follows in relation to the objectives set out in Chapter 1.

- One objective was to: "*To develop a new image de-noising procedure to reduce noise in solar images using a wavelet transform-based method, to improve the peak to signal noise ratio (PSNR) and Image quality assessment parameters such as MSE , SSIM , SC, NCC, and NAE are carried out for comparing performance"*

This goal was achieved in the third chapter, where a system for the automatic detection of sunspots was proposed. The system entails the development of a variety of algorithms. Firstly, a method for enhancing image quality and resolution was proposed. This method used a combination of discrete and stationary wavelet decomposition for five wavelengths before using the discrete wavelet transform, which has been proven to give a superior performance compared with existing image enhancement techniques. Its excellent performance was confirmed by the results of a

thorough evaluation, which used visual observation, as well as the calculation of the MSE and PNSR values.

In order to improve the visibility of solar features, such as sunspots, in images at five wavelengths, a combination of discrete and stationary wavelet decomposition was used for the enhancement of resolution in SODISM images. The practical implementation is described, and an evaluation has been presented regarding the effectiveness of the segmentation of SODISM features using algorithms and focusing on sunspots. Additionally, a substantial volume of available SODISM data (around 900 images) was used to compare the performance of the proposed technique with images at different wavelengths, and other interpolation techniques. It is clear from Figures 30-39, that the images enhanced by the proposed technique are sharper and cleaner than those from other methods.

Furthermore, the quantitative comparisons of the PSNR and MSE also confirm the effectiveness of the proposed method in enhancing resolutions. In all cases, the proposed technique shows good performances in terms of PSNR, MSE, and visual clarity. A further test was carried out using the proposed enhancement technique, where the sunspot regions in the benchmark, enhanced and tested images at 393 nm and 535 nm wavelengths were all segmented in an equivalent manner. A numerical comparison of the segmentation results was also applied. Using the enhanced images for segmentation, and calculating the resulting FFs represented a further test of the proposed technique, which was applied to the sunspot features identified in 393 nm and 535 nm images from SODISM. The results showed that an improvement is possible by first enhancing the test images using the proposed technique. Moreover, from inspecting the overall results, the FF assessment shows that the enhanced images are closer to the benchmark images than the test images. This represents improvement in enhancement images while keeping the original details.

- A further two objectives were to: "*Develop a new algorithm to detect solar features in five different SODISM wavelengths and calculate their filling factor*" and to "*Provide visual maps for the segmented solar features.*"

These objectives were also achieved in the third chapter and the FFs were discussed in the fourth chapter. The new segmentation method improves the automatic detection of sunspots in solar images, and the methods used in this thesis provide two different approaches to detect segmented sunspots in a solar image. The first method, which was applied to the FITS images at wavelengths 535 nm, 607 nm and 782 nm, can be used for automatic sunspot detection in full solar disc images. The second method can be used after the automatic image cleaning procedures for the elimination of limb darkening and noise removal. A combination of algorithms was used to provide a new image enhancement approach, and many filters were used to reduce the effect of noise, artefacts, and unwanted features. Automated sunspot detection saves time by reducing the need for analysis, and providing a more accurate identification of sunspots as a consequence. Also, a novel image segmentation pipeline has been proposed, developed and integrated into the system. Accordingly, the two proposed algorithms combined with morphological and image differencing operations, form part of a completely automatic sunspot detection system. A thorough evaluation of each component of the detection system was conducted with multiple images. Two different image formats for the SODISM data acquired from PICARD were used.

This work has also shown that, despite degradation in the quality of SODISM images, the automatic detection of sunspots is possible with suitable pre-processing procedures. Moreover, the proposal of a novel method to summarise results a significant contribution to knowledge. The cataloguing procedure presented in Chapter 4, gives a clear representation of the analysis results of the SODISM data.

Comparisons have been conducted between the results from the automated segmentation using algorithm 3 with data from the NOAA catalogue, and the overall sunspot recognition rate is approximately 98%, as shown in Table 7. These results provide excellent matches with the NOAA catalogue, and provide strong supporting evidence for the efficacy of the proposed automatic method. Despite success in four wavelengths, the only method with weaknesses in the segmentation was at 215 nm; this may be

because they are very noisy images and have a short operating time. As mentioned in Figure 41, channel 215 lost more than 90% of the normalized intensity, and this may have caused the degradation. Finally, another weakness concerning the FF is the lack of any other available source for comparison, except with SOHO.

Based on results and despite the low resolution qualities of some wavelength images, the methods achieve the goals set, and provide clearly segmented maps, and assessment by visual inspection or by mathematical calculation illustrates the efficiency of the methods. In other words, the evaluation of the performance of this method, specific to solar application, demonstrates clear visual results.

- The third objective for this study was to: "*Provide comparisons between SODISM images and SOHO images for the same time period*" and to "*Provide a Filling Factor catalogue for SODISM images"*

This was achieved in the fourth chapter in which the FF catalogues of SODISM images were calculated, and their accuracy was investigated through comparisons with the SOHO images for the same period. At 535 nm, the correlation coefficient between the two datasets was found to be 98.5 %, while at 607 nm it was 99% for the SOHO images. This may be related to the fact that the SOHO data corresponds to a different wavelength (676.8 nm) from that of the SODISM images. However, this is the first automated method to achieve≈0.99 corrections between SOHO and SODISM in the FF ratio. Moreover, the statistical calculations in Table 15 provide measurements for skewedness, and indicate that the data is normally distributed. Therefore, the parametric scores can be relied upon, providing confidence that the method applied to the SODISM data achieved a mean sunspot region of 1.876×10^{3}, and a corresponding average sunspot ratio of 6.62×10^{-4}. Furthermore, it was deduced that any image picked from the dataset could be expected to have a FF signature of ±3 standard deviations, and the value of σ can be conveniently retrieved from Table 16. Finally, the IQR-based outlier detection conducted has revealed that seven images within the data lie outside the normal distribution. Hence, like any other real-

world data, it is expected that these few outliers (≈0.032% of the dataset) can exhibit a few irregularities. Generally, the previously defined metrics present researchers with a new method of summarising hundreds of experiments in a more concise manner. To show the detection result, the comparison has been made between the SODISM catalogue and the NOAA catalogue in the following categories; number of sunspots detected their location and size, their time of capture. The accuracy comparison for period August 2010 to January 2014 shows a correlation coefficient of 95.17% Figure 97. Moreover, the sunspot sizes are matching 93% in the NOAA catalogue (Figure 98). Overall, the statistical means of characterising the SODISM data presented here are new to the research community. Therefore, the FF catalogue can be used as a template to conduct comparisons. Should researchers embark on work similar to that presented in this thesis, it is hoped they will utilise this procedure and present their results for comparative analysis.

- The final objective of this research was to: "*Provide evaluation techniques to assess the accuracy of segmentation*".

The evaluation methods used have met this study's aims and objectives. Three techniques, explained in the fifth chapter, were adopted to achieve this goal. The first technique involved a visual assessment, which involved labelling the segmented image in red or white, and superimposing it on the original image. This method was helpful in determining features, but not fully accurate, because some of the smaller sunspot detection could be missed due to the lack of visibility and human error. Indeed, tiny sunspots are difficult to identify by the human eye (Colak and Qahwaji, 2007). The second assessment technique supported the first, and was achieved by comparing the results with other reliable results; this meant a comparison with the NOAA catalogue, where the calculated corrections were 98% (Table 7). Moreover, over the lifetime of SODISM this was 95.17%. This assessment suggests that the automated segmentation approach successfully achieved the goals.

The third assessment concerned the FF calculations, which adopted two methods. Firstly, a comparison of the SOHO FF was calculated using

different methods over the same time period with SODISM FFs, and at three different wavelengths, namely 535 nm, 607 nm and 782 nm. The resulting correlation coefficients were 98.5%, 99%, and 93%, respectively. These results reflect the strong correlation between the SOHO FF results and the SODISM FF results. The best wavelength at which to automate segmentation to detect sunspots is 607 nm, and the second best is 535 nm. Figures 75 and 78 show W.L 535 and 607 compared with the SOHO FF ratio, but this is most evident over the period from 22th September 2010 to 29th September 2010, when there is a somewhat lower correlation coefficient (0.95 for 607 nm). Nevertheless, this is still better than those from wavelength 535 nm for the same period, which shows a correlation of 0.92. By inspecting SODISM images over this period, it was shown that some images are brighter than the rest and this might affect the segmentation results.

Secondly, a statistical method, shown in Table 16, was used; more specifically, the parametric measures employed were three statistical tests, namely mean, variance and skewness. Moreover, the hand median and inter-quartile ranges (IQR) were used as the non-parametric measures. Parametric and non-parametric statistics (Sheskin, 2003) were used to describe the data. This has been discussed in the previous section and is considered a novel method to describe SODISM FF results because this is the first study to adopt it for such a purpose.

6.1.3 General conclusion

The main achievements of this thesis are the development of a method to enhance SODISM images and the provision of an automatic method for SODISM image segmentation. However, the particular contribution of this research is that it is probably the first study to develop a SODISM image FF calculation, and to compile catalogues for the results. This thesis addresses the problem of enhancing and segmenting SODISM images using novel methods. These methods combine the sound practice of tried and tested approaches and the innovative approaches associated with new technology, thus exploiting a combination of advantages to achieve the goals of the study to overcome the challenges associated with SODISM images, including poor

quality and the lack of a comparative reference. These problems were overcome using a new statistical method as well as FF, and by a comparison with reliable results in SOHO. The most important lesson, and therefore knowledge gained, was that accurate data pre-processing enhanced the segmentation results, and therefore achieved the thesis aims. Despite these successes, some aspects remain for future research due to a lack of time. A further limitation was the enhancement for wavelength 215 nm, which for this study's method does not provide acceptable results. This could be due to many factors, such as the presence of a lot of noise and the darkening at this wavelength; as such, these images need further experimentation and analysis to successfully extract sunspot maps from this wavelength. In addition, the SODISM instrument's point spread function (PSF) effect, which uses a small sunspot (less than a pixel) after applying the enhancement method, and SODISM images, may be corrected. This can be addressed by, first estimating the point-spread-function of the blurring process, and then by applying an inverse de-blurring transformation. Moreover, it is possible to quantify the gain with this new technique. Indeed, the resolution enhancement for SODISM images is a promising future area of investigation because this is still a pure research area.

The comparison and data from this research could encourage new research into SODISM images as the scientific applications of this study have contributed valid information for a FF catalogue and presented a database for future work. However, further investigation is required to study the evolution of FF over a longer period, as required with any satellite. Thus, the limitations of this work will be addressed in the future work section as potential areas of development.

6.2 Future work

SODISM images remain a potential area for further study, as this is the first thesis in this area, to the best of the researcher's knowledge, although there are many incomplete works. Therefore, for future research, the simplest extension of the work presented here would be to continue the study in its current form. Different methods have been presented for segmentation and enhancement, which are supported by the results and

assessment methods; however, there are still many different adaptations, experiments, and tests that have not been investigated due to a lack of time and the huge volumes of data produced by the PICARD satellite that could provide a target for further scientific research, The research presented in this thesis can be further extended by considering the following suggestions:

- The enhancement method presented in Chapter 3, can be developed to improve the resolution image pixel for SODISM images. This could be achieved through adopting enhancement operations, (such as the Fourier transform) for the SODISM images and by conducting a comparison with the results.
- A comparison could be conducted between the FF of SODISM images over the lifetime of a satellite and other FF calculations from SDO (should more calculations become available in the future). From this, it could be possible to develop the method by separating the umbra and penumbra in order to extract information on the umbra of the spot, and to obtain more information on the formation and evolution of sunspots.
- Another new area for future work could be the identification of a unified system for detecting and tracking active regions (AR) on the solar disk in SODISM images, which accords with NOAA detection and calculation properties, and includes magnetic and chromospheric properties. This could enable an investigation into the relationship between the properties of other solar activities, such as Coronal Mass Ejections (CME), and the evolution of these properties in relation to flares, flare classes, and flare properties.
- Furthermore, bright region detection could be investigated in future work, whilst the design of automated systems for detecting bright regions, calculating their FF and comparing them with other satellites, such as SDO, could also be considered. It is worth noting that, at this point in time, there is no comparative research for this study. The limitation with ghosts in images, which impacts on the effectiveness of this enhancement method, is a further area of future research. Such work could isolate the ghost from the image before applying pre-processing approaches; this could lead to the resolution of problems in, and

development of, a method for SODISM that helps to extract a bright regions map.

Further investigation is required to analyse SODISM images and a number of sunspot properties in relation to different flare intensities. This might mean utilising other technologies, such as feature selection or machine learning algorithms, or investigating data extracted from other satellites, such as SDO images, to extract further findings and confirm the knowledge gained from this study over a longer period. This could improve the catalogue and add FF sunspots in far more detail than previously available.

6.3 Research resources

In this research, the following resources have been adopted:

- C++ and MATLAB programming languages were adopted to implement the IMC model, in Chapter 3. The programs calculated the NAE, C S, NCC MSIM, MSN and PSNR. An executable program was included in the CD attached to this thesis.

- C++ programming language was adopted to implement the association algorithms, in Chapter 3 and 4. The program associated the FF.

References

A. Temizel and T. Vlachos (2005) ‘Wavelet domain image resolution enhancement using cycle-spinning’, IET ELECTRONICS LETTERS, 41(3). doi: 10.1049/el.

Abd, Mehmmood A, Sarab M and V. Zharkova V. (2010) ‘Technological Developments in Networking, Education and Automation’. doi: 10.1007/978-90-481-9151-2.

Adipranata R., Budhi G. and Setiahadi B. (2013) ‘Automatic classification of sunspot groups for space weather analysis’, International Journal of Multimedia and Ubiquitous Engineering, 8(3), pp. 41–54.

Aeronautics N. (2011) ‘Dwarf Stars’, NASA Goddard Space Flight Center and the Space Telescope Science Institute. Available at: www.hubblesite.org/hubble_discoveries/science_year_in_review.

Alasta A., Algamudi A., Qahwaji R., Ipson S., and Nagerm T., (2017) ‘Automatic sunspots detection on SODISM solar images’, in 2017 Seventh International Conference on Innovative Computing Technology (INTECH). IEEE, pp. 115–119. doi: 10.1109/INTECH.2017.8102429.

Alomari M. H., (2009) ‘Engineering System Design for Automated Space Weather Forecast’.

Alptekin Temizel (2007) ‘IMAGE RESOLUTION ENHANCEMENT USING WAVELET DOMAIN HIDDEN MARKOV Alptekin Temizel Graduate School of Informatics , METU , Ankara , Turkey’, Proc. Int. Conf. Image Process, 5, p. v 381-384.

Ashamari O., Qahwaji R., Ipson S., and Scholl M., (2015) ‘Identification of photospheric activity features from SOHO/MDI data using the ASAP tool’, (c). doi: 10.1051/swsc/2015013.

Ashamari O., Qahwaji R., Ipson S., and Schöll M. (2015) ‘Identification of photospheric activity features from SOHO/MDI data using the ASAP tool’, Journal of Space Weather and Space Climate, 5, p. A15. doi: 10.1051/swsc/2015013.

Assus P., Irbah A, Bourget P., and Corbard T., (2008) ‘Monitoring the scale factor of the PICARD SODISM instrument’, Astronomische

Nachrichten, 329(5), pp. 517–520. doi: 10.1002/asna.200710987.

Atkins C., Bouman C. and Allebach J. (2001) ‘OPTIMAL IMAGE SCALING USING PIXEL CLASSIFICATION Hewlett-Packard Laboratories 1501 PageMill Rd.,M/S 4U-6 Palo Alto, CA 94304’, Proc. ICIP, 3, pp. 864–867.

Bagawade Ramdas, B. K. S. and P. P. M. (2012) ‘Wavelet Transform Techniques for Image Resolution Enhancement : A Study’, International Journal of Emerging Technology and Advanced Engineering, 2(4), pp. 167–172.

Bosco, A. et al. (2003) ‘A Noise Reduction Filter For Full-Frame Data Imaging Devices’, IEEE Transactions on Consumer Electronics, 49(3), pp. 676–682. doi: 10.1109/TCE.2003.1233800.

Boyat, A. K. and Joshi, B. K. (2015) ‘a R Eview P Aper : N Oise M Odels in D Igital I Mage P Rocessing’, Signal & Image Processing, 6(2), pp. 63–75. doi: 10.5121/sipij.2015.6206.

Bratsolis, E. and Sigelle, M. (1998) ‘Solar image segmentation by use of mean field fast annealing’, Astronomy and astrophysics, Supplement Series, 131(August), pp. 371–375.

Canfield, R. C. (2005) ‘Solar Active Regions’, Encyclopedia of Astronomy and Astrophysics, (785998), pp. 1–6. doi: 10.1888/0333750888/2023.

Chan, T. F. and Zhou, H. M. (2003) ‘ENO-Wavelet Transforms for Piecewise Smooth Functions’, SIAM Journal on Numerical Analysis, 40(4), pp. 1369–1404. doi: 10.1137/S0036142900370915.

Christopoulou, Eugenia B., A. N. S. and A. A. G. (2002) ‘Trous” wavelet traksform wrsus ci.asical methods for the improvement of solar images’, 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No.02TH8628), Cat. No.02, pp. 885–888.

Colak, T. and Qahwaji, R. (2007) ‘Automatic sunspot classification for real-time forecasting of solar activities’, Proceedings of the 3rd International Conference on Recent Advances in Space Technologies, RAST 2007, (May), pp. 733–738. doi: 10.1109/RAST.2007.4284089.

Corbard T., Salabert D., Boumier P., Appourchaux T., Hauchecorne A., Journoud P., N. A. Gelly, B., Hochedez, J. F., Irbah, A., Meftah M.,

Meisonnier M., Nyeki S., Pfiffner D., Roth H., Rozanov E., Spescha M., Wehrli C., Werner L., and Wyss J. (2013) 'The PREMOS/PICARD instrument calibration', Journal of Physics Conference Series, p. 440. doi: 10.1088/0026-1394/46/4/S13.

Corbard T., Boumier P., Appourchaux T., Jim´enez-Reyes J., B.Gelly B., and the PICARD team. (2008) 'Helioseismology program for the PICARD satellite', Astronomische Nachrichten, 329(May), pp. 508–516. doi: 10.1002/asna.200710986.

Curto, J. J., Blanca, M. and Martínez, E. (2008) 'Automatic sunspots detection on full-disk solar images using mathematical morphology', Solar Physics, 250(2), pp. 411–429. doi: 10.1007/s11207-008-9224-6.

D. C. Stenning, T. C. M. Lee, D. A. Van Dyk, V. Kashyap, J.Sandell, C. A. Y. (2013) 'Morphological feature extraction for statistical learning with applications to solar image data', Statistical Analysis and Data Mining, 6(4), pp. 329–345.

Daneels, D., van C., David N., Carlton W., Equitz W., Barber R., and Fierens F. (1993) 'Interactive outlining: an improved approach using active contours', in Niblack, C. W. (ed.). International Society for Optics and Photonics, pp. 226–233. doi: 10.1117/12.143652.

Davey A., Bogart R., Gurman J., Hill (NSO), Hourclé J., Martens P., I. Suarez-Sola I., Tian K., and Yoshimura K. (2004) 'Implementation of the Virtual Solar Observatory', Bulletin of the American Astronomical Society, Vol. 36 (January), pp. 796.

Demirel H., and Anbarjafari, G. (2010) 'Satellite image resolution enhancement using complex wavelet transform', IEEE Geoscience and Remote Sensing Letters, 7(1), pp. 123–126. doi: 10.1109/LGRS.2009.2028440.

Demirel, H. and Anbarjafari, G. (2011a) 'Discrete wavelet transform-based satellite image resolution enhancement', IEEE Transactions on Geoscience and Remote Sensing, 49(6 PART 1), pp. 1997–2004. doi: 10.1109/TGRS.2010.2100401.

Demirel, H. and Anbarjafari, G. (2011b) 'Discrete Wavelet Transform-Based Satellite Image Resolution Enhancement', 49(6), pp. 1997–2004.

Eddy, J. a. (2010) ‘Sun, the Earth, and Near-Earth Space: A Guide to the Sun-Earth System’.

Etcheto P., Meftah M., Meissonnier M., Irbah A., Assus, P. and Thuillier, G. (2011) ‘Stray light in PICARD SODISM instrument: design, check, flight results, and alignment issues’, (October), p. 81310K. doi: 10.1117/12.896434.

Fernandez Borda, R. A., Mininni P. D., Mandini C. H., Gómez D., Bauer O., and Rovira M. (2002) ‘Automatic solar flare detection using neural network techniques’, Solar Physics, 206, pp. 347 – 357. doi: 10.1023/A.

Field A. (2013) Discovering statistics using IBM SPSS statistics. Sage.

Foukal P. and Wiley InterScience (Online service) (2004) Solar astrophysics. Wiley-VCH. Available at: https://books.google.co.uk/books?hl=en&lr=&id=z73VCHWjxjkC&oi=fnd&pg=PP2&dq=Foukal,+P.+V.,+Solar+astrophysics,+Wiley-VCH,+2008.&ots=vsbZlKmlzh&sig=vV-OfMJk9lrh67nAg1l15ChYG5g#v=onepage&q&f=false (Accessed: 16 March 2018).

Fowler J. E. (2005) ‘The Redundant Discrete Wavelet Transform and Additive Noise’, IEEE Signal Processing Letters, 12(9), pp. 629–632. doi: 10.1109/LSP.2005.853048.

G. Pradels, T. Guinle, G. Thuillier, A. Irbah, J-P. Marcovici, C. Dufour, D. Moreau, C. Noel, M. Dominique, T. Corbard, M. Hadjara, S. Mekaoui, C. W. (no date) ‘The PICARD Payload Data Centre’, American Institute of Aeronautics and Astronautics, pp. 1–17.

Gabhel M. and Hiradhar A. (2014) ‘Comparative Analysis of Various Image Denoising Techniques : A Review Paper’, 3(7), pp. 1877–1881.

‘Galileo’s Sunspot Letters’ (no date). The British Library. Available at: https://www.bl.uk/collection-items/galileos-sunspot-letters (Accessed: 23 December 2018).

González-Audícana M., X. OTAZU, O. FORS and A. SECO Department. (2005) ‘Comparison between Mallat’s and the “à trous” discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images’, International Journal of Remote Sensing, 26(3),

pp. 595–614. doi: 10.1080/01431160512331314056.

Gonzalez R. C. (2002) Digital Image Processing, Leonardo. doi: 10.2307/1574313.

Gravetter F. J. and Wallnau, L. B. (2016) Statistics for the behavioral sciences. Cengage Learning.

Gray L. J. et al. (2010) 'SOLAR INFLUENCES ON CLIMATE', Reviews of Geophysics, 48(4), p. RG4001. doi: 10.1029/2009RG000282.

Hardie R. C. and Barner, K. E. (1994) 'Filters for Signal Restoration', 3(2), pp. 192–206.

Harikrishna O. and Maheshwari, a (2012) 'Satellite Image Resolution Enhancement using DWT Technique', (5), pp. 274–278.

Hathaway D. H., (2015) 'The solar cycle', Living Reviews in Solar Physics, 12(1), pp. 1–81. doi: 10.1007/lrsp-2015-4.

Heinz N. and Dietrich L. (1994) 'solar limb darkening', Hamburger Sternwarte, Germany, pp. 91–114. Available at: https://link.springer.com/article/10.1007%2FBF00712494.

Hu Q. and Sonnerup B. U. (2002) 'Reconstruction of magnetic clouds in the solar wind: Orientations and configurations', Journal of Geophysical Research: Space Physics, 107(A7), pp. 1–15. doi: 10.1029/2001JA000293.

Hubrecht J. B. (1913) 'Some problems of Asironomy', (36), pp. 398 – 401.

Izenman A. J. (1983) 'J. R. Wolf and H. A. Wolfer: An Historical Note on the Zurich Sunspot Relative Numbers', Journal of the Royal Statistical Society. Series A (General). WileyRoyal Statistical Society, 146(3), p. 311. doi: 10.2307/2981658.

Kamboj P. and Rani V. (2013) 'A Brief Study of Various Noise Model and Filtering Techniques', Journal of Global Research in Computer Science, 4(4), pp. 166–171.

Kaur B. and Shukla M. (2014) 'Image De-noising and its Methods : A Survey', 3(6), pp. 234–238.

Khaire M. G. and Shelkikar R. P. (2013) 'Resolution Enhancement of images with interpolation and DWT–SWT Wavelet domain components', International Journal of Application or Innovation in Engineering and

Management, 2(9), pp. 242–8. Available at: http://ijaiem.org/volume2issue9/IJAIEM-2013-09-26-061.pdf%5Cnhttp://www.ijaiem.org/volume2issue9/IJAIEM-2013-09-26-061.pdf.

Khan G. B. (2016) 'A Modified Non Linear Median Filter for the Removal of Medium Density Random Valued Impulse Noise', 147(8), pp. 17–22.

Khidse S. (2014) 'Implementation and Comparison of Image Enhancement Techniques', 96(4), pp. 9–16.

Khobragade P, B. and Thakare, S., S. (2014) 'Image Compression Techniques : A Review', (IJCSIT) International Journal of Computer Science and Information Technologies, 5(1), pp. 272–275.

Kimmel R. and Bruckstein a. M. (2003) 'Regularized Laplacian zero crossings as optimal edge integrators', International Journal of Computer Vision, 53(3), pp. 225–243. doi: 10.1023/A:1023030907417.

Kumar B. S., Nagaraj S. and Tech M. (2013a) 'Discrete and Stationary Wavelet Decomposition for IMAGE Resolution Enhancement', International Journal of Engineering Trends and Technology (IJETT) - Volume4 Issue7- July 2013 Discrete, 4(July), pp. 2885–2889.

Kumar B. S., Nagaraj S. and Tech M. (2013b) 'Discrete and Stationary Wavelet Decomposition for IMAGE Resolution Enhancement', 4(July), pp. 2885–2889.

Kumar V. (2017) 'Impacts of space weather on aviation', CAWCR Technical Report No. 081, p. 110. doi: 10.1002/2014WR015483.

Kumar V. and Gupta P. (2012) 'Importance of Statistical Measures in Digital Image Processing', International Journal of Emerging Technology and Advanced Engineering, 2(8), pp. 56–62. doi: doi=10.1.1.366.2595.

Kuriakose J. and Joy J. (2014) 'Image Fusion Using Kuwahara Filter', 5(4), pp. 5944–5949.

Lakshmi M. S. D. (2013) 'Robust Satellite Image Resolution Enhancement Based On Interpolation Of Stationary Wavelet Transform', 4(6), pp. 1365–1369.

Leavline E. J., Asir D. and Singh A. G. (2013) 'Salt and Pepper Noise Detection and Removal in Gray Scale Images: An Experimental

Analysis', International Journal of Signal Processing Image Processing and Pattern Recognition, 6(5), pp. 343–352. doi: 10.14257/ijsip.2013.6.5.30.

Lefebvre S. and Rozelot J. P. (2004) 'A new method to detect active features at the solar limb', Solar Physics. Kluwer Academic Publishers, 219(1), pp. 25–37. doi: 10.1023/B:SOLA.0000021818.97402.1e.

Li, X. and Orchard M. T., (2001) 'New edge-directed interpolation', IEEE Transactions on Image Processing, 10(10), pp. 1521–1527. doi: 10.1109/83.951537.

Lysaker M., Lundervold A. and Tai X., (2003) 'Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time', IEEE Transactions on image processing. IEEE, 12(12), pp. 1579–1590.

Al-Omari M., Qahwaji R., Colak T.and Ipson. S. (2010) 'Machine Leaning-Based Investigation of the Associations between CMEs and Filaments', Solar Physics, 262(2), pp. 511–539. doi: 10.1093/bjsw/bcs140.

Meftaha M., Meissonniera M., Irbaha A., Abbakia S., Assusb P., Bertrana E., Duboisc J.P., Ducourta E., Dufoura C., Marcovicia J.P., Poieta G., A. J. V. and T. Gally (no date) 'The Space instrument SODISM and the Ground instrument SODISM II', Proceedings of SPIE - The International Society for Optical Engineering, 7731, pp. 1–12. doi: 10.1117/12.851227.

McIntosh, P. S. (1990) 'The classification of sunspot groups', Solar Physics. Kluwer Academic Publishers, 125(2), pp. 251–267. doi: 10.1007/BF00158405.

Meftah M., Meissonniera M., Irbaha A., Abbakia S., Assusb P., Bertrana E., Duboisc J.P., Ducourta E., Dufoura C., Marcovicia J.P., Poieta G., Vieaua A.J. and Thuilliera G. (2010) 'The space instrument SODISM and the ground instrument SODISM II', in Oschmann, Jr., J. M., Clampin, M. C., and MacEwen, H. A. (eds) Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave. Edited by Oschmann, Jacobus M., Jr.; Clampin, Mark C.; MacEwen, Howard A. Proceedings of the SPIE, Volume 7731, article id. 773145, 12 pp.

(2010)., p. 773145. doi: 10.1117/12.851227.

Meftah M., Irbah A., Le Letty R., Bataille A., Ducourt E., Poiet G. and Privat M., (2011) ‘The picard/sodism pointing mechanism: from the design to the flight performances’, (September), pp. 28–30.

Meftah, M., Irbaha A., Hauchecornea A., and Hochedez J. (2013) ‘PICARD payload thermal control system and general impact of the space environment on astronomical observations’, (May 2013), p. 87390B. doi: 10.1117/12.2010178.

Meftah M., Hochedez J., Irbah A., Hauchecorne A., Boumier P.,Corbard T., S. Turck-Chi`eze S., Assus P., Bertran E., Bourget P. Buisson F., Chaigneau M., Dam´e L., Djafer D., Dufour C., Etcheto P., Ferrero P., Hers´e M., Marcovici J., Meissonnier M., Morand F., Poiet G., Prado J., Renaud C., Rouanet N., Rouz´e M., Salabert D., and Vieau A. (2014a) ‘Picard SODISM, a Space Telescope to Study the Sun from the Middle Ultraviolet to the Near Infrared’, Solar Physics, 289(3), pp. 1043–1076. doi: 10.1007/s11207-013-0373-x.

Meftah, M. Hochedez J., Irbah A., Hauchecorne A., Boumier P.,Corbard T., S. Turck-Chi`eze S., Assus P., Bertran E., Bourget P. Buisson F., Chaigneau M., Dam´e L., Djafer D., Dufour C., Etcheto P., Ferrero P., Hers´e M., Marcovici J., Meissonnier M., Morand F., Poiet G., Prado J., Renaud C., Rouanet N., Rouz´e M., Salabert D., and Vieau A. (2014b) ‘Picard SODISM, a Space Telescope to Study the Sun from the Middle Ultraviolet to the Near Infrared’, Solar Physics, 289(3), pp. 1043–1076. doi: 10.1007/s11207-013-0373-x.

Meftah M., Hochedez J., Irbah A., Hauchecorne A., Boumier P.,Corbard T., S. Turck-Chi`eze S., Assus P., Bertran E., Bourget P. Buisson F., Chaigneau M., Dam´e L., Djafer D., Dufour C., Etcheto P., Ferrero P., Hers´e M., Marcovici J., Meissonnier M., Morand F., Poiet G., Prado J., Renaud C., Rouanet N., Rouz´e M., Salabert D., and Vieau A. (2014) ‘PICARD SODISM, a space telescope to study the Sun from the middle ultraviolet to the near infrared’, (January), pp. 1–38.

Meftah M., Hauchecorne A., A. Irbah A., Corbard T.,2, Ikhlef R., Morand F., C. Renaud, Riguet F., and Pradal F. (2015) ‘on the Constancy of the

Diameter of the Sun During the Rising Phase of Solar Cycle 24', The Astrophysical Journal, 808(1), p. 4. doi: 10.1088/0004-637X/808/1/4.

Meftah, M., Irbah, A., Hauchecorne A., Corbard T., Turck-Chièze S., Hochedez J., Boumier P., Chevalier A., Dewitte S., Mekaoui S. and Salabert D., (2015) 'On the Determination and Constancy of the Solar Oblateness', Solar Physics, 290(3), pp. 673–687. doi: 10.1007/s11207-015-0655-6.

Meftah, M. Corbard T., A. Hauchecorne A., Irbah A.,; Boumier P.,; A. Chevalier A.,; Schmutz W., Ikhlef R., F. Morand F., C. Renaud C., Hochedez J.,; G. Cessateur G., Turck-Chièze S., Salabert D., M. Rouzé M., van Ruymbeke M., Zhu P.,; Kholikov S.,; Koller S., Conscience C., Dewitte S., Damé L.,and D. Djafer D.,. (2016) 'Main results of the PICARD mission', (July), p. 99040Z. doi: 10.1117/12.2232027.

Meftah, M. Corbard T., Hauchecorne A., Morand F., Ikhlef R., Chauvineau B., Renaud C., Sarkissian A., and Damé L. (2018) 'Solar radius determined from PICARD/SODISM observations and extremely weak wavelength dependence in the visible and the near-infrared', Astronomy & Astrophysics. EDP Sciences, 616, p. A64. doi: 10.1051/0004-6361/201732159.

Memon, F., Unar M. A. L. I. and Memon S. (2015) 'Image Quality Assessment for Performance Evaluation of Focus Measure Operators', 34(4), pp. 379–386.

Milbourne T. W., Haywood R. D., Phillips D. F., Saar S. H., Cegla H. M., Cameron A. C. and Walsworth, R. L. (2019) 'HARPS-N Solar Radial-Velocity Variations Are Dominated By Large, Bright Magnetic Regions', 874. doi: 10.3847/1538-4357/ab064a.

Mishra J. and Mitra, S. (2014) 'Image Denoising using Brute Force Thresholding Algorithm', 3(9), pp. 830–836.

Mrudula D. and Lakshmi G. N. (2014) 'Joint DWT , SWT based MRI image nhancement technique and denosing based on', IPASJ International Journal of Electronics & Communication (IIJEC), 2(9), pp. 64–69.

N.R.C. Committee on the societal and economic impacts of and SpaceWeather (2009) Severe Space Weather Events--Understanding

Societal and Economic Impacts. doi: 10.17226/12643.

Narayana, B. R. V. S. and Nirmala, K. (2012) 'Image Resolution Enhancement by Using Stationary and Discrete Wavelet Decomposition', International Journal of Image, Graphics and Signal Processing, 4(11), pp. 41–46. doi: 10.5815/ijigsp.2012.11.06.

NASA (2004) 'Solar wind', pp. 1–2. Available at: https://nasasearch.nasa.gov/search?query=solar+winds&affiliate=nasa&utf8=✓.

National Aeronautics and Space Administration (2017) 'TSIS-1 TRACKING THE POWER SUPPLY TO', pp. 1–4. Available at: https://www.nasa.gov/sites/default/files/atoms/files/tsis_brochure-dec8.pdf.

Pambrun J. F. and Noumeir R. (2015) 'Computed Tomography Image Compressibility and Limitations of Compression Ratio-Based Guidelines', Journal of Digital Imaging, 28(6), pp. 636–645. doi: 10.1007/s10278-015-9791-7.

Piao Y. P. Y., Shin I. S. I. and Park H. W. P. H. W. (2007) 'Image Resolution Enhancement using Inter-Subband Correlation in Wavelet Domain', 2007 IEEE International Conference on Image Processing, 1, pp. 445–448. doi: 10.1109/ICIP.2007.4378987.

Plunkett S. P. (2005) 'The Extreme Solar Storms of October to November 2003 : an overview', (November 2003), pp. 91–98.

Pratt W. K. (2000) 'Image Segmentation', Computer Vision, pp. 579–622. doi: 10.1002/9780470097434.ch17.

Pucha R., Hiremath K. M. and Gurumath S. R. (2016) 'Development of a Code to Analyze the Solar White-Light Images from the Kodaikanal Observatory : Detection of Sunspots , Computation of Heliographic Coordinates and Area'. doi: 10.1007/s12036-016-9370-4.

Qahwaji R. and Colak T. (2005a) 'Automatic detection and verification of solar features', International Journal of Imaging Systems and Technology, 15(0), pp. 199–210. doi: 10.1002/ima.20053.

Qahwaji R. and Colak T. (2005b) 'Automatic detection and verification of solar features', International Journal of Imaging Systems and

Technology, 15, pp. 199–210. doi: 10.1002/ima.20053.

Rahvar S. (2015) ‘Gravitational microlensing I: A unique astrophysical tool’, International Journal of Modern Physics D, 24(07), p. 1530020. doi: 10.1142/S0218271815300207.

Raju M. N., Babu G. R., and Kumar, K. A. (2014) ‘Image Resolution Enhancement by Using Multi Resolution Transform’, IJCSIET--International Journal of Computer Science information and Engg., Technologies, 3(4), pp. 2–6.

Qahwaji R.,, Ipson S., Ashamari O. (2015) ‘SOLID D3.5 Filling factors catalogue for PICARD images SOLID’, First European Comprehensive SOLar Irradiance Data exploitation, pp. 1–8.

Rao Y. R., Prathapani N., and Nagabhooshanam E. (2014) ‘Application of Normalized Cross Correlation To Image Registration’, International Journal of Research in Engineering and Technology, 3(5), pp. 12–16. Available at: http://www.ijret.org.

Rousseeuw P. J. and Hubert M. (2011) ‘Robust statistics for outlier detection’, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. Wiley Online Library, 1(1), pp. 73–79.

Rouzé, M. Hauchecorne A., .F. Hochedez J., Irbah A., Meftah M.,. Corbard T., Turck-Chièze S. Boumier P., Dewitte S. and Schmutz W. (2014) ‘The PICARD Scientific Mission: status of the program’, in SpaceOps 2014 Conference. Reston, Virginia: American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2014-1705.

Sangeetha S., and Krishna, Y. H. (2013) ‘Image Resolution Enhancement Technique Based on the Interpolation of the High Frequency Subbands Obtained by DWT’, 4(July), pp. 3061–3067.

Schmutz, W. Fehlmann A., Hülsen G., Meindl P., Winkler R., Thuillier G., Blattner P., Buisson F., Egorova T., Finsterle W., Fox N., Gröbner J., Hochedez J., Koller S., Meftah M., Mireille Meisonnier M., Nyeki S., Pfiffner D., Roth H., Rozanov E., Spescha M., Wehrli C., Werner L. and Wyss J. (2009) ‘The PREMOS/PICARD instrument calibration’, Metrologia, 46(4), pp. S202–S206. doi: 10.1088/0026-1394/46/4/S13.

Sciacchitano F. , Dong Y. ,; and Hansen P. C. (2017) ‘Image reconstruction

under non-Gaussian noise', Citation. Available at: www.compute.dtu.dk.

Sheskin, D. J. (2003) Handbook of parametric and nonparametric statistical procedures. crc Press.

Shmulevich, I. and Arce, G. R. (2001) 'Admitting Negative Weights', Signal Processing, 8(12), pp. 313–316.

Singh, P. (2016) 'A Comparative Study to Noise Models and Image Restoration Techniques', 149(1), pp. 18–27.

Solanki, S. K. (2003) 'Sunspots: An overview', Astronomy and Astrophysics Review, 11(2–3), pp. 153–286. doi: 10.1007/s00159-003-0018-4.

Srinivasan, G. N. (2008) 'Segmentation Techniques for ATDR', WSEAS Transactions on Computers, 2(2), pp. 75–81. Available at: http://www.wseas.us/e-library/transactions/computers/2008/27-729.pdf.

Colak T. and Qahwaji R. (2007) 'Automated Prediction of Solar Flares Using Neural Networks and Sunspots Associations', 96(Jule). doi: 10.1007/978-3-642-20505-7.

Thompson W. T. (2006) 'Coordinate systems for solar image data', Astronomy & Astrophysics, 449(2), pp. 791–803. doi: 10.1051/0004-6361:20054262.

Trochim W. M. K. and Donnelly, J. P. (2001) 'Research methods knowledge base'. Atomic Dog Pub.

Turmon M., Pap J. and Mukhtar S. (2002) 'Statistical pattern recognition for labeling solar active regions: application to SOHO/MDI imagery', The Astrophysical Journal, 568, pp. 396–407. doi: 10.1086/338681.

Valens, C. (1999) 'A Really Friendly Guite to Wavelets'. Available at: https://www.cs.unm.edu/~williams/cs530/arfgtw.pdf.

Varnan, C. S. (2011) 'Image Quality Assessment Techniques in Spatial Domain', International Journal of Computer Science and Technology, 2(3), pp. 177–184. Available at: http://www.ijcst.com/vol23/1/sasivarnan.pdf%0Ahttp://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.219.5535.

Verma J., and Agrawal P. (2013) 'A Survey of Linear and Non-Linear Filters for Noise Reduction', International Journal of Advance Research in Computer Science and Management Studies, 1(3), pp. 18–25. doi:

10.1111/j.1475-4762.2011.01069.x.

Verma, R. and Ali, J. (2013) ‘A Comparative Study of Various Types of Image Noise and Efficient Noise Removal Techniques’, International Journal of Advance Research in Computer Science and Software Engineering, 3(10), pp. 617–622.

Waldmeier M., (1976) ‘The sunspot-activity in the years 1961 - 1975.’, Astron. Mitt. Eidgen. Sternw. Zürich, Nr. 346, 19 p., 346. Available at: http://adsabs.harvard.edu/abs/1976MitSZ.346.....W (Accessed: 23 December 2018).

Young I. T., Gerbrands J. J. and van Vliet L. J. (2007) ‘Fundamentals of Image Processing (v.2.3)’, (November 2012), pp. 1–112. doi: 10.1007/978-3-319-46316-2_11.

Yu L., Deng L. and Feng S. (2014) ‘Automated sunspot detection using morphological reconstruction and adaptive region growing techniques’, Proceedings of the 33rd Chinese Control Conference, CCC 2014, 33(0), pp. 7168–7172. doi: 10.1109/ChiCC.2014.6896184.

Zharkov S., Zharkova V., Ipson S. (2009) ‘Statistical properties of sunspots in 1996 – 2004: I. Detection, North – South asymmetry and area distribution. Solar Phys’, 228(1), pp. 377–397. doi: 10.1007/s11207.

Zharkova V.,Ipson S., Aboudarham J., and Bentley B.. (2002) ‘Survey of Image Processing Techniques’, (December).

Zharkova V., Ipson S., Benkhalil A., and Zharkov S. (2005) ‘Feature recognition in solar images’, Artificial Intelligence Review, 23(3), pp. 209–266. doi: 10.1007/s10462-004-4104-4.